华为职业认证通过者权益

通过**任一项**华为职业认证，您即可在华为在线学习网站(*http://learning.huawei.com/cn*) 享有如下特权：

- **1、华为E-learning 课程学习**
  - **内容**：**所有华为职业认证E-Learning课程**，扩展您在其他技术领域的技术知识
  - **方式**：请提交您的 “华为账号” 和注册账号的 “email地址” 到 *Learning@huawei.com 申请权限。*
- **2、华为培训教材下载**
  - **内容**：**华为职业认证培训教材+华为产品技术培训教材，**覆盖企业网络、存储、安全等诸多领域
  - **方式**：登录*华为在线学习网站*，进入 “*华为培训->面授培训*” ，在具体课程页面即可下载教材。
- **3、华为在线公开课(LVC)优先参与**
  - **内容**：企业网络、UC&C、安全、存储等诸多领域的职业认证课程，华为讲师授课，开班人数有限
  - **方式**：开班计划及参与方式请详见LVC排期：
    http://support.huawei.com/learning/NavigationAction!createNavi#navi[id]=_16
- **4、学习工具 eNSP**
  - *eNSP (Enterprise Network Simulation Platform)*, 是由华为提供的免费的、可扩展的、图形化网络仿真工具。主要对企业网路由器和交换机进行硬件模拟，完美呈现真实设备实景；同时也支持大型网络模拟，让大家在没有真实设备的情况下也能够进行实验测试。
- 另外, 华为建立了知识分享平台 华为认证论坛。您可以在线与华为技术专家交流技术，与其他考生分享考试经验，一起学习华为产品技术。（http://support.huawei.com/ecommunity/bbs/list_2247.html）



华为认证系列教程

HCNA-HNTD入门
华为网络技术与设备



华为技术有限公司

版权声明

华为认证系列教程

HCNA-HNTD华为网络技术与设备

第2.0版本

华为认证体系介绍

依托华为公司雄厚的技术实力和专业的培训体系，华为认证考虑到不同客户对ICT技术不同层次的需求，致力于为客户提供实战性、专业化的技术认证。

根据ICT技术的特点和客户不同层次的需求，华为认证为客户提供面向十二个方向的三级认证体系。

HCNA主要面向IP网络维护工程师，以及其他希望学习IP网络知识的人士。HCNA认证在内容上涵盖TCP/IP基础、路由、交换等IP网络通用基础知识以及华为数据通信产品、通用路由平台VRP特点和基本维护。

HCNP-R&S主要面向企业级网络维护工程师、网络设计工程师以及希望系统深入地掌握路由、交换、网络调整及优化技术的人士。HCNP-R&S包括IESN（Implementing Enterprise Switching Networks，部署企业级交换网络）、IERN（Implementing Enterprise Routing Networks，部署企业级路由网络）、IENP（Improving Enterprise Network Performance，提升企业级网络性能）三个部分。内容上涵盖IPv4路由技术原理深入以及在VRP中的实现；交换技术原理深入以及在VRP中的实现；网络安全技术、高可靠性技术和Qos技术等高级IP网络技术以及在华为产品中的实现。

HCIE-R&S旨在培养能够熟练掌握各种IP网络技术，精通华为产品的维护、诊断和故障排除；具备大型IP网络规划、设计和优化的IP网络大师。

华为认证协助您打开行业之窗，开启改变之门，屹立在ICT世界的潮头浪尖！

建议进阶关系（虚线箭头）　必备进阶关系（实线箭头）

ICT

	路由交换		无线局域网	无线	传送网	安全	统一通信		视讯	云计算	存储	ICT融合设计
工程师级	HCNA(HCDA)											
工程师级			HCNA-WLAN	HCNA-LTE	HCNA-Transmission	HCNA-Security	HCNA-UC	HCNA-CC	HCNA-VC	HCNA-Cloud	HCNA-Storage	HCNA-Design
资深工程师级	HCNP-R&S (HCDP)	HCNP-Carrier (HCDP-Carrier)	HCNP-WLAN	HCNP-LTE	HCNP-Transmission	HCNP-Security	HCNP-UC	HCNP-CC	HCNP-VC	HCNP-Cloud	HCNP-Storage	HCNP-Design
专家级	HCIE-R&S	HCIE-Carrier	HCIE-WLAN	HCIE-LTE	HCIE-Transmission	HCIE-Security	HCIE-UC	HCIE-CC	HCIE-VC	HCIE-Cloud	HCIE-Storage	HCIE-Design
架构师级	HCAr											

前言

简介

本书为HCNA认证培训教程，专门适用于准备参加HCNA考试的学员。对于希望通过在华为VRP平台上进行实际操作和演练，从而加强认识和理解数据通信原理的在校学生和专业人员，本书也极具参考价值。

内容描述

本书共包含五个Module，全面地介绍了构建一个基本的IP网络所涉及的各种主要技术，重点描述了交换、路由和网络服务等基础内容，以及这些内容是如何在VRP上配置和实现的。

Module 1系统地介绍了TCP/IP协议模型，侧重讲述了数据链路层、网络层和传输层的功能和作用，其主要目的是帮助读者加深对数据通信中“层次”的理解，并且熟悉和掌握数据在网络中的端到端传输过程。

Module 2介绍了华为通用路由平台VRP的基础知识及其操作指导，主要包含了VRP的基本结构和特性、VRP的命令行基础、VRP的文件系统、以及如何管理VRP等内容。

Module 3介绍了以太网交换机的基本工作原理，并对局域网中广泛使用的STP/RSTP技术进行了较为详细的描述和分析。

Module 4介绍了路由的基本原理以及在VRP中的配置实现，内容包含IP路由基

础、静态路由、以及IGP中最为常用的RIP路由协议和OSPF路由协议。

Module 5介绍了三种常见的网络应用：DHCP，FTP和Telnet，以及这些应用在VRP上的配置和实现。

本书的目的在于：结合华为的产品实现，引导读者从基础开始，循序渐进地熟悉和掌握以路由交换技术为核心的数据通信知识。对于基础知识较为薄弱的读者，建议严格按照本书的编排内容进行顺序地阅读学习；其他读者可根据自身的情况进行有选择性地阅读学习。

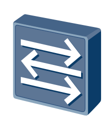
读者知识背景

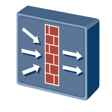
本课程为华为认证 HCNA v2.0 入门课程，要求读者具有最基本的网络知识背景。

本书常用图标


路由器


汇聚层交换机


接入层交换机


防火墙


AP


PC


便携电脑


普通服务器


电话


IP电话

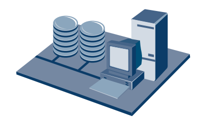

DSLAM


RADIUS服务器


邮件服务器


存储服务器

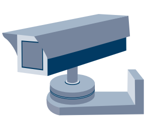

应用服务器

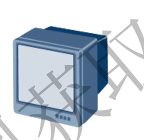

NMS


智能手机


PAD


球机


带云台的枪机


IP电视


AC

目录

Module-0

企业网络基础


企业网络架构介绍
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

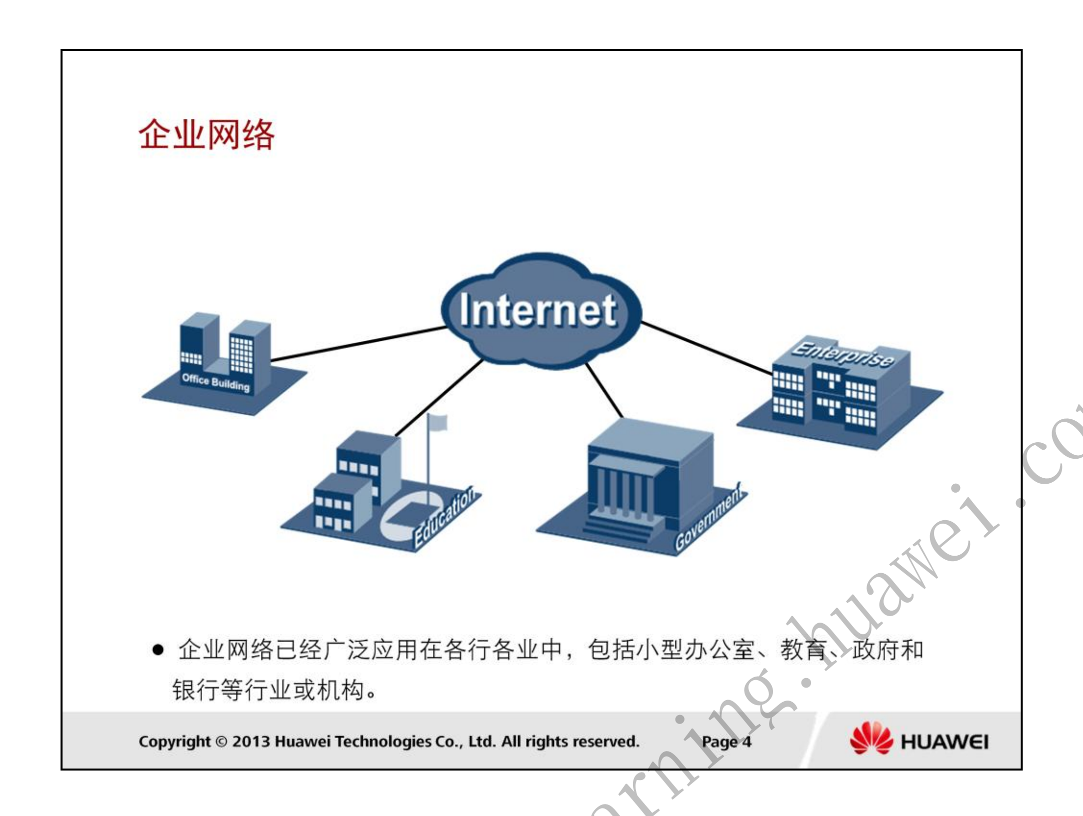
企业的业务总是在不断地发展，对网络的需求也是在不断地变化，这就要求企业网络应该具备适应这种需求不断变化的能力。因此，了解企业网络的架构是如何适应业务的需求将变得十分必要。



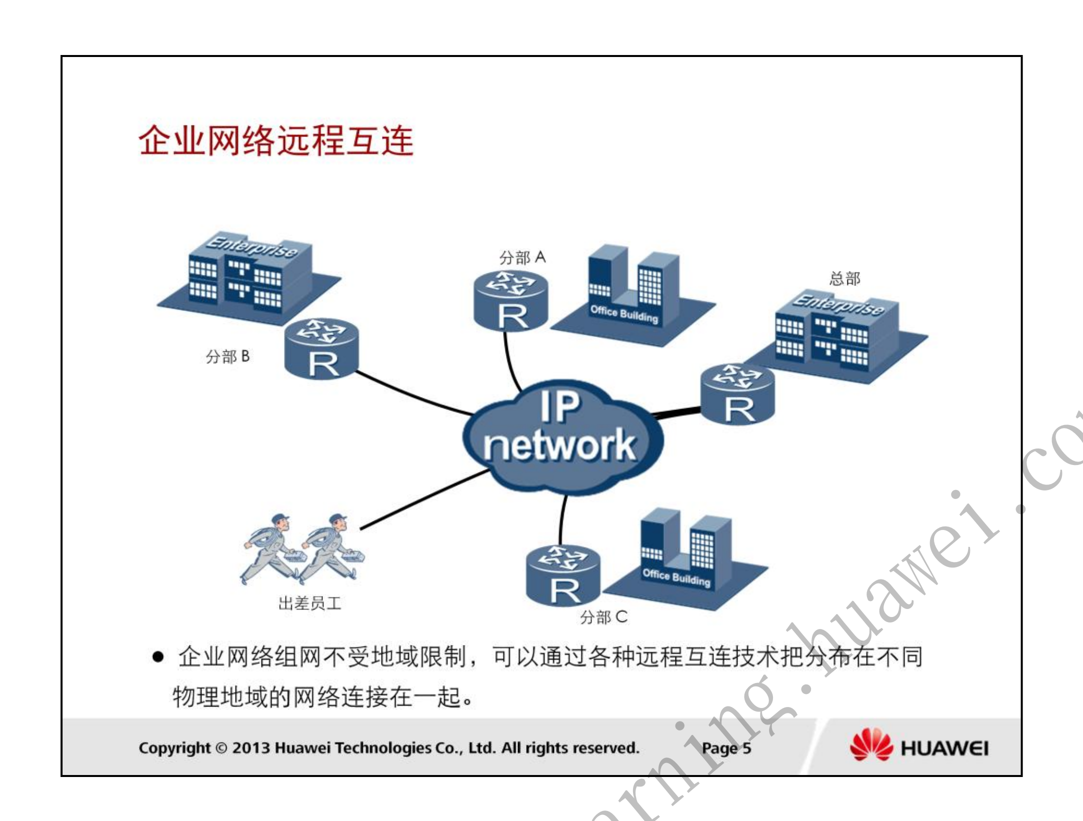


学完本课程后，您应该能：

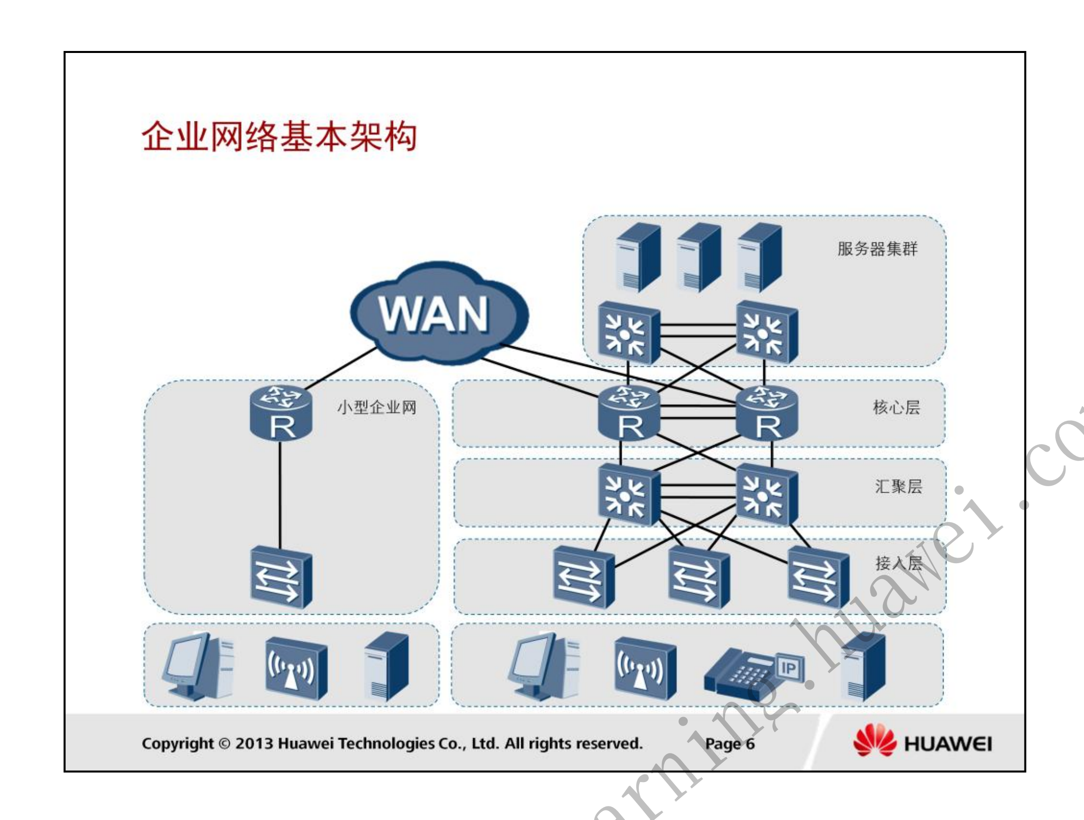
- 了解企业网络的基本架构



最初，企业网络是指某个组织或机构的网络互联系统。企业使用该互联系统主要用于共享打印机、文件服务器等，使用email实现用户间的高效协同工作。现在，企业网络已经广泛应用在各行各业中，包括小型办公室、教育、政府和银行等行业或机构。



大型企业的网络往往跨越了多个物理区域，所以需要使用远程互连技术来连接企业总部和分支机构，从而使得出差的员工能随时随地接入企业网络实现移动办公，企业的合作伙伴和客户也能够及时高效的访问到企业的相应资源及工具。在实现远程互连的同时，企业还会基于对数据的私密性和安全性的考虑对远程互连技术进行选择。



企业网络架构很大程度上取决于企业或机构的业务需求。小型企业通常只有一个办公地点，一般采用扁平网络架构进行组网。这种扁平网络能够满足用户对资源访问的需求，并具有较强的灵活性，同时又能大大减少部署和维护成本。小型企业网络通常缺少冗余机制，可靠性不高，容易发生业务中断。

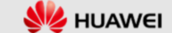
大型企业网络对业务的连续性要求很高，所以通常会通过网络冗余备份来保证网络的可用性和稳定性，从而保障企业的日常业务运营。大型企业网络也会对业务资源的访问进行控制，所以通常会采用多层网络架构来优化流量分布，并应用各种策略进行流量管理和资源访问控制。多层网络设计也可以使网络易于扩展。大型企业网络采用模块化设计能够有效实现网络隔离并简化网络维护，避免某一区域产生的故障影响到整个网络。



总结

- 小型企业网络和大型企业网络的组网有什么差别?
- 大型企业网络设计的基本思想是什么?

Page 7



1. 小型企业网络通常采用扁平网络结构，网络扩展能力低。如需支持未来不断增长的用户，应采用多层网络结构。大型企业网络用户较多，通常采用层次化结构以支持网络的扩展和用户的增长。
2. 在设计大型企业网络时必须首先考虑企业业务的特点，在保证网络性能满足业务需求的前提下，必须还要考虑网络的可用性、稳定性、可扩展性、安全性和可管理性，以保证企业业务的正常运营和发展。


谢谢
www.huawei.com

Module-1

简单的IP网络


传输介质简介
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

通信网络除了包含通信设备本身之外，还包含连接这些设备的传输介质，如同轴电缆、双绞线和光纤等。不同的传输介质具有不同的特性，这些特性直接影响到通信的诸多方面，如线路编码方式、传输速度和传输距离等。

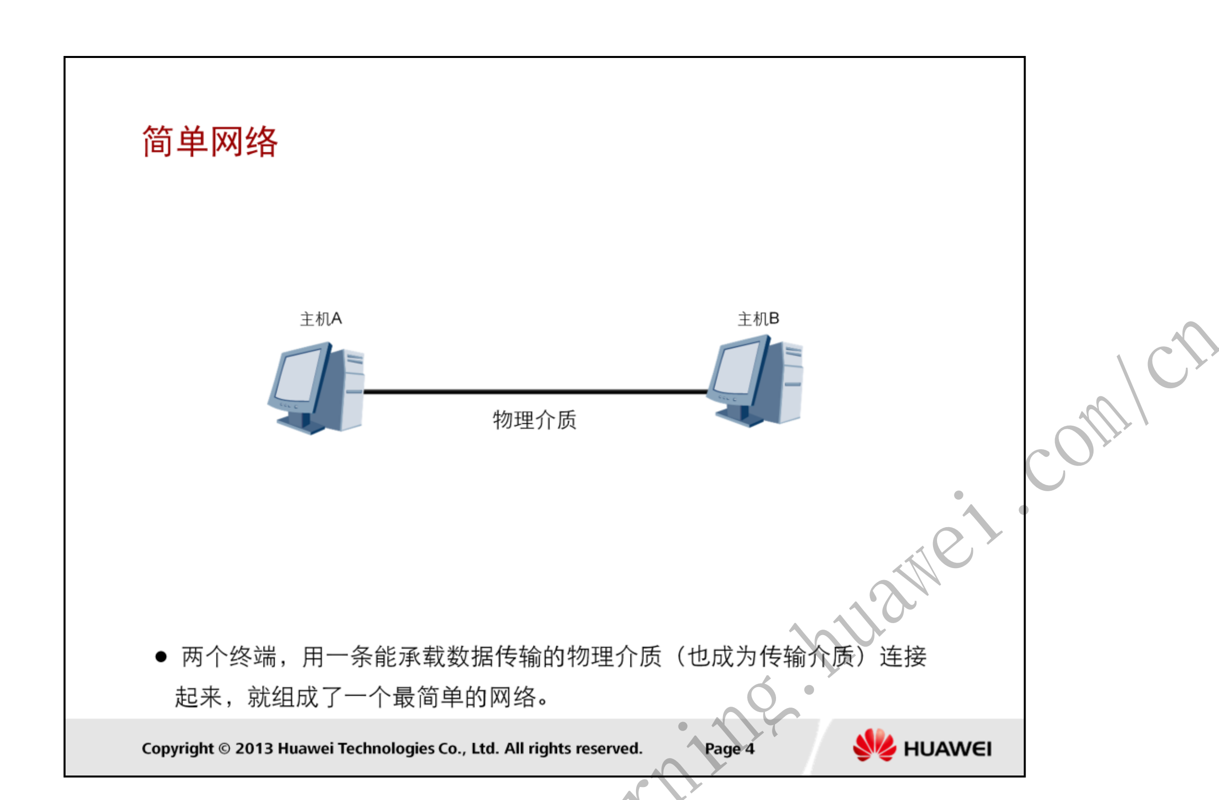




学习目标

学完本课程后，您应该能：

- 了解一些常见的传输介质
- 理解冲突域和双工模式的基本概念



终端相互传递信息和资源共享的需求是网络产生的主要原因。

终端可以产生、发送和接收数据，网络是终端建立通信的媒介，终端通过网络建立连接。用来传输数据的载体称为介质，网络可以使用各种介质进行数据传输，包括物理线缆，无线电波等。

网络就是通过介质把终端互连而成的一个规模大、功能强的系统，从而使得众多的终端可以方便地互相传递信息，共享信息资源。



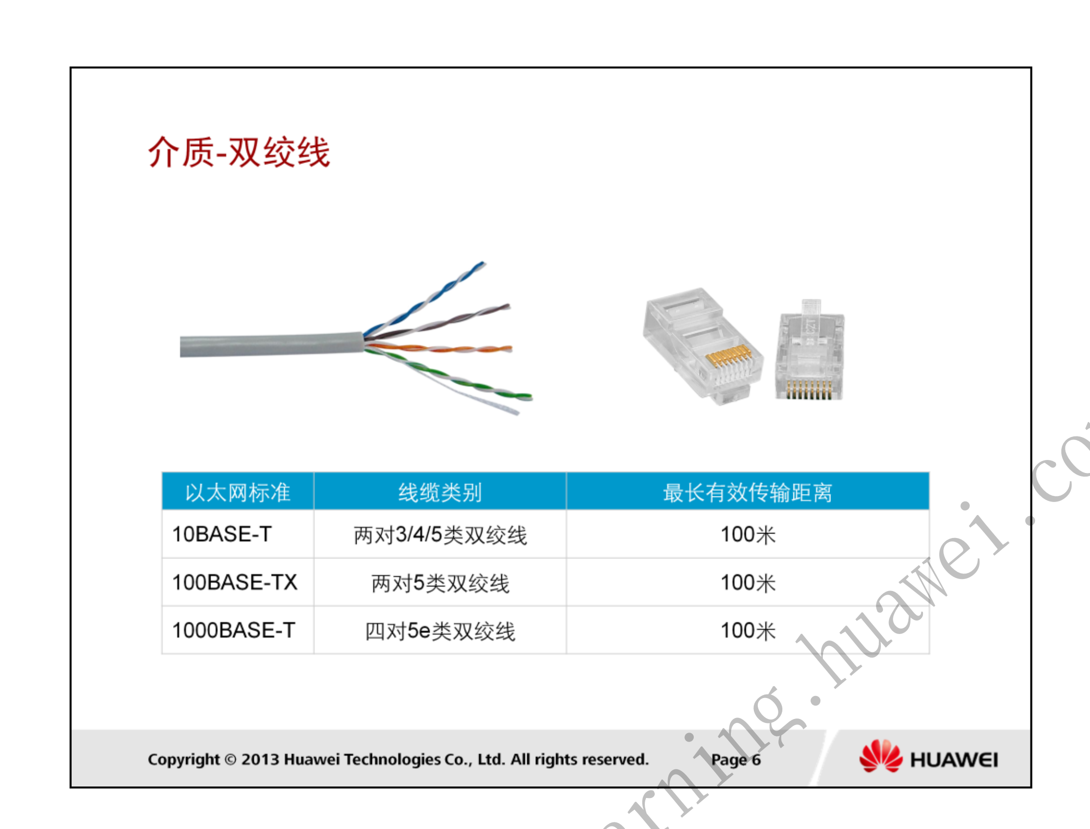
以太网标准	电缆类别	最长有效传输距离
10BASE5	粗同轴电缆	500米
10BASE2	细同轴电缆	185米

同轴电缆是一种早期使用的传输介质，同轴电缆的标准分为两种，10BASE2和10BASE5。这两种标准都支持10Mbps的传输速率，最长传输距离分别为185米和500米。一般情况下，10Base2同轴电缆使用BNC接头，10Base5同轴电缆使用N型接头。

现在，10Mbps的传输速率早已不能满足目前企业网络需求，因此同轴电缆在目前企业网络中很少应用。

10BASE5和10BASE2是早期的两种以太网标准，它们均采用同轴电缆作为传输介质。10BASE5和10BASE2所使用的同轴电缆的直径分贝为9.5mm和5mm，所以前者又称为粗缆，后者又称为细缆。一般情况下，10BASE5的同轴电缆使用N型接头，10BASE2的同轴电缆使用BNC接头。10BASE5和10BASE2都支持10Mbps的传输速率，最长有效传输距离分别是500米和185米。

目前，这两种以太网已基本被淘汰，企业网中也几乎不再使用它们。



与同轴电缆相比双绞线（Twisted Pair）具有更低的制造和部署成本，因此在企业网络中被广泛应用。双绞线可分为屏蔽双绞线(Shielded Twisted Pair，STP)和非屏蔽双绞线(Unshielded Twisted Pair，UTP)。屏蔽双绞线在双绞线与外层绝缘封套之间有一个金属屏蔽层，可以屏蔽电磁干扰。双绞线有很多种类型，不同类型的双绞线所支持的传输速率一般也不相同。例如，3类双绞线支持10Mbps传输速率；5类双绞线支持100Mbps传输速率，满足快速以太网标准；超5类双绞线及更高级别的双绞线支持千兆以太网传输。双绞线使用RJ-45接头连接网络设备。为保证终端能够正确收发数据，RJ-45接头中的针脚必须按照一定的线序排列。

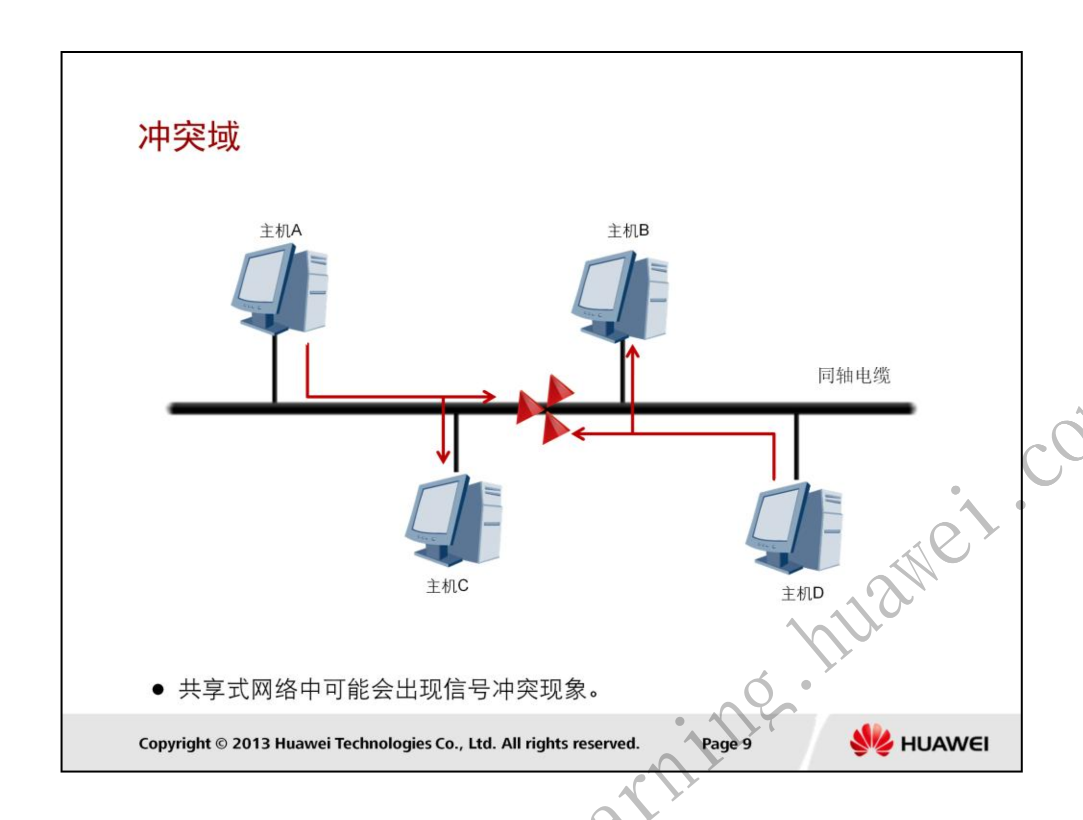


以太网标准	线缆类别	最长有效传输距离
10BASE-F	单模/多模光纤	2000 米
100BASE-FX	单模/多模光纤	2000 米
1000BASE-LX	单模/多模光纤	316 米
1000BASE-SX	多模光纤	316 米

双绞线和同轴电缆传输数据时使用的是电信号，而光纤传输数据时使用的是光信号。光纤支持的传输速率包括10Mbps，100Mbps，1Gbps，10Gbps，甚至更高。根据光纤传输光信号模式的不同，光纤又可分为单模光纤和多模光纤。单模光纤只能传输一种模式的光，不存在模间色散，因此适用于长距离高速传输。多模光纤允许不同模式的光在一根光纤上传输，由于模间色散较大而导致信号脉冲展宽严重，因此多模光纤主要用于局域网中的短距离传输。光纤连接器种类很多，常用的连接器包括ST，FC，SC，LC连接器。



网络通信中常常会用到各种各样的串口电缆。常用的串口电缆标准为RS-232，同时也是推荐的标准。但是RS-232的传输速率有限，传输距离仅为6米。其他的串口电缆标准可以支持更长的传输距离，例如RS-422和RS-485的传输距离可达1200米。RS-422和RS-485串口电缆通常使用V.35接头，这种接头在上世纪80年代已经淘汰，但是现在仍在帧中继、ATM等传统网络上使用。V.24是RS-232标准的欧洲版。RS-232本身没有定义接头标准，常用的接头类型为DB-9和DB-25。现在，RS-232已逐渐被FireWire、USB等新标准取代，新产品和新设备已普遍使用USB标准。

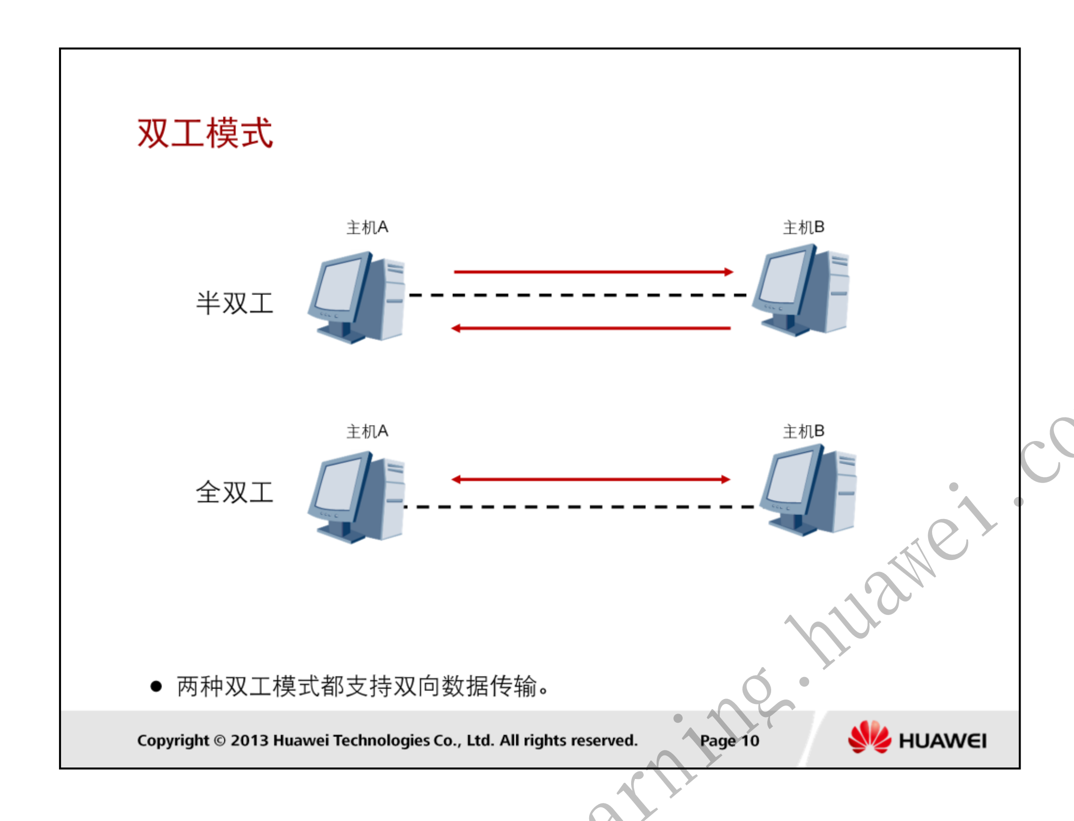


如图是一个10BASE5以太网，每个主机都是用同一根同轴电缆来与其它主机进行通信，因此，这里的同轴电缆又被称为共享介质，相应的网络被称为共享介质网络，或简称为共享式网络。共享式网络中，不同的主机同时发送数据时，就会产生信号冲突的问题，解决这一问题的方法一般是采用载波侦听多路访问/冲突检测技术（Carrier Sense Multiple Access/Collision Detection）。

CSMA/CD的基本工作过程如下:

1. 终端设备不停地检测共享线路的状态。如果线路空闲，则可以发送数据；如果线路不空闲，则等待一段时间后继续检测（延时时间由退避算法决定）。
2. 如果有另外一个设备同时发送数据，两个设备发送的数据会产生冲突。
3. 终端设备检测到冲突之后，马上停止发送自己的数据，并发送特殊阻塞信息，以强化冲突信号，使线路上其他站点能够尽早检测到冲突。
4. 终端设备检测到冲突后，等待一段时间之后再进行数据发送（延时时间由退避算法决定）。

CSMA/CD的工作原理可简单总结为：先听后发，边发边听，冲突停发，随机延迟后重发。



半双工：在半双工模式（half-duplex mode）下，通信双方都能发送和接收数据，但不能同时进行。当一台设备发送时，另一台只能接收，反之亦然。对讲机是半双工系统的典型例子。

全双工：在全双工模式（full-duplex mode）下，通信双方都能同时接收和发送数据。电话网络是典型的全双工例子。

以太网上的通信模式包括半双工和全双工两种：

半双工模式下，共享物理介质的通信双方必须采用CSMA/CD机制来避免冲突。例如，10BASE5以太网的通信模式就必须是半双工模式。

全双工模式下，通信双方可以同时实现双向通信，这种模式不会产生冲突，因此不需要使用CSMA/CD机制。例如，10BASE-T以太网的通信模式就可以是全双工模式。

同一物理链路上相连的两台设备的双工模式必须保持一致。



总结

- 企业网络中部署千兆以太网时使用哪种传输介质?
- 什么是冲突域?
- CSMA/CD的作用是什么?



1. 千兆以太网传输必须使用超5类标准及以上的双绞线，或者使用千兆及更高等级的光纤。
2. 冲突域是一个通过共享物理介质进行双向传输的所有节点的集合。当同一冲突域中的主机同时发送数据时，数据到达目的地之前可能会发生冲突。
3. CSMA/CD是一种在共享式网络上检测并避免冲突的机制。CSMA/CD的工作过程如下:

   1）终端设备不停地检测共享线路的状态。如果线路空闲，则可以发送数据；如果线路不空闲，则一直等待。

   2) 如果有另外一个设备同时发送数据，两个设备发送的 数据必然产生冲突，导致线路上的信号混乱。

   3) 终端设备检测到这种冲突之后，马上停止发送自己的数据。

   4) 终端设备发送一连串干扰脉冲，然后等待一段时间之后再进行数据发送。


谢谢
www.huawei.com


以太网帧结构
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

网络中传输数据时需要定义并遵循一些标准，以太网是根据IEEE 802.3标准来管理和控制数据帧的。了解IEEE802.3标准是充分理解以太网中链路层通信的基础。

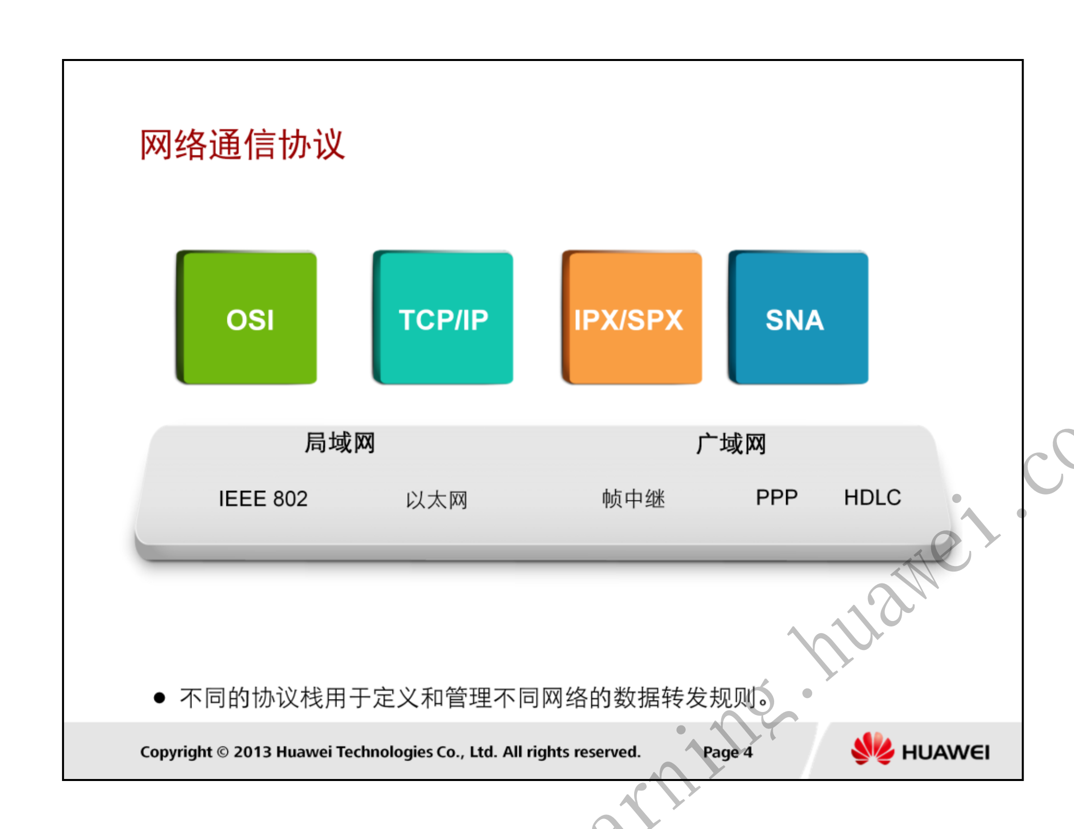




学习目标

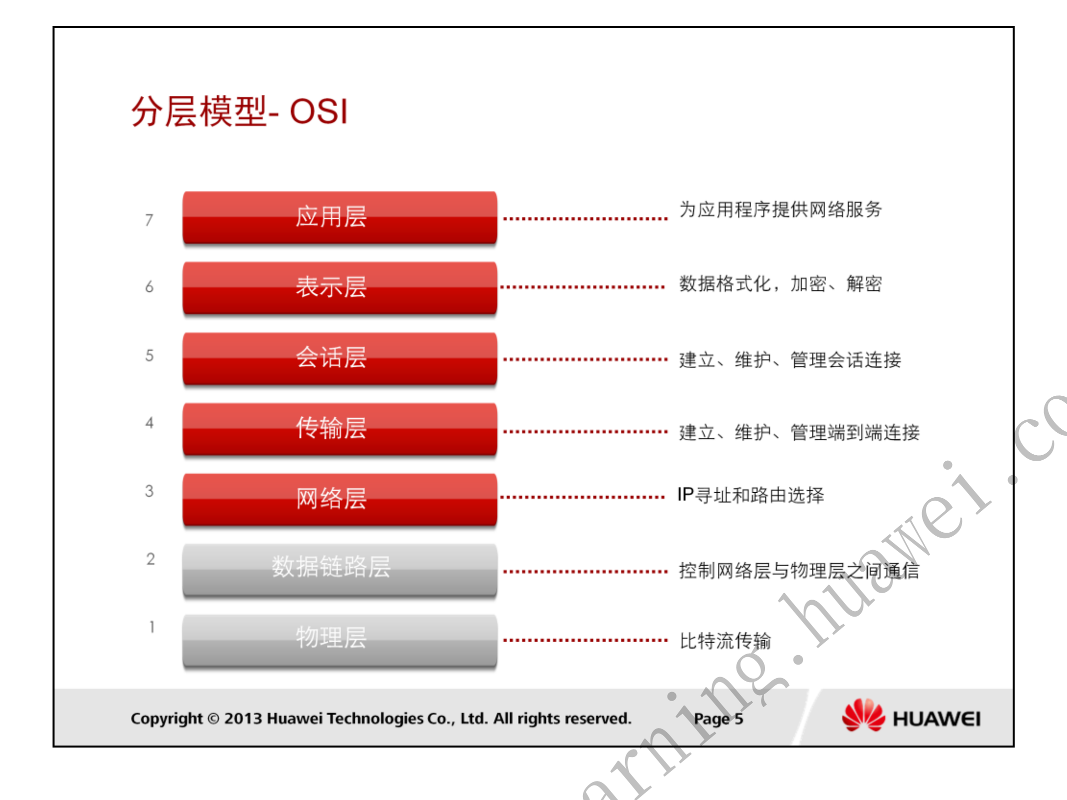
学完本课程后，您应该能：

- 理解分层模型的作用
- 掌握以太网中数据帧的结构
- 掌握MAC地址的作用
- 掌握以太网中数据帧转发的过程



20世纪60年代以来，计算机网络得到了飞速发展。各大厂商和标准组织为了在数据通信网络领域占据主导地位，纷纷推出了各自的网络架构体系和标准，如IBM公司的SNA协议，Novell公司的IPX/SPX协议，以及广泛流行的OSI参考模型和TCP/IP协议。同时，各大厂商根据这些协议生产出了不同的硬件和软件。标准组织和厂商的共同努力促进了网络技术的快速发展和网络设备种类的迅速增长。

网络通信中，"协议"和"标准"这两个词汇常常可以混用。同时，协议或标准本身又常常具有层次的特点。一般地，关注于逻辑数据关系的协议通常被称为上层协议，而关注于物理数据流的协议通常被称为低层协议。IEEE 802就是一套用来管理物理数据流在局域网中传输的标准，包括在局域网中传输物理数据的802.3以太网标准。还有一些用来管理物理数据流在使用串行介质的广域网中传输的标准，如帧中继FR（Frame Relay），高级数据链路控制HDLC（High-Level Data Link Control），异步传输模式ATM（Asynchronous Transfer Mode）。



国际标准化组织ISO于1984年提出了OSI RM（Open System Interconnection Reference Model，开放系统互连参考模型）。OSI参考模型很快成为了计算机网络通信的基础模型。

OSI参考模型具有以下优点：简化了相关的网络操作；提供了不同厂商之间的兼容性；促进了标准化工作；结构上进行了分层；易于学习和操作。

OSI参考模型各个层次的基本功能如下：

物理层**:** 在设备之间传输比特流，规定了电平、速度和电缆针脚。

数据链路层：将比特组合成字节，再将字节组合成帧，使用链路层地址（以太网使用MAC地址）来访问介质，并进行差错检测。

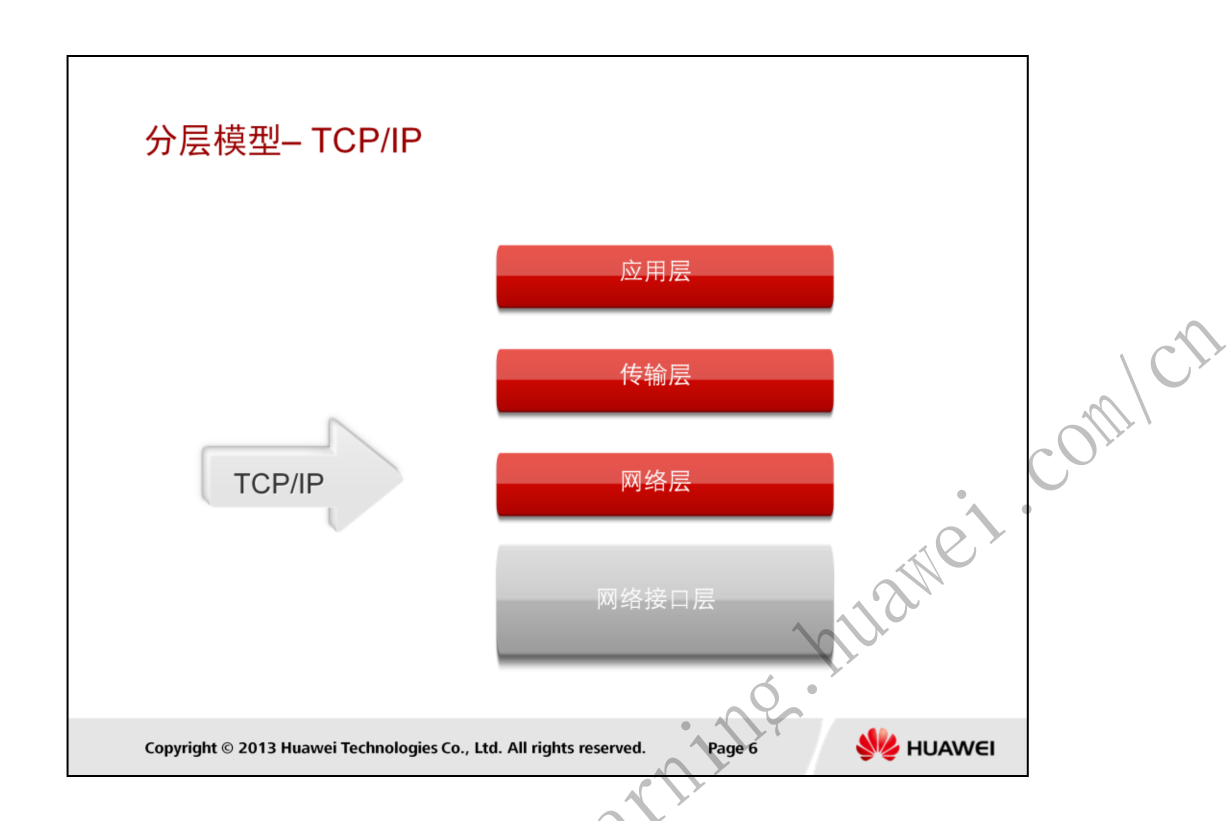
网络层：提供逻辑地址，供路由器确定路径。

传输层：提供面向连接或非面向连接的数据传递以及进行重传前的差错检测。

会话层：负责建立、管理和终止表示层实体之间的通信会话。该层的通信由不同设备中的应用程序之间的服务请求和响应组成。

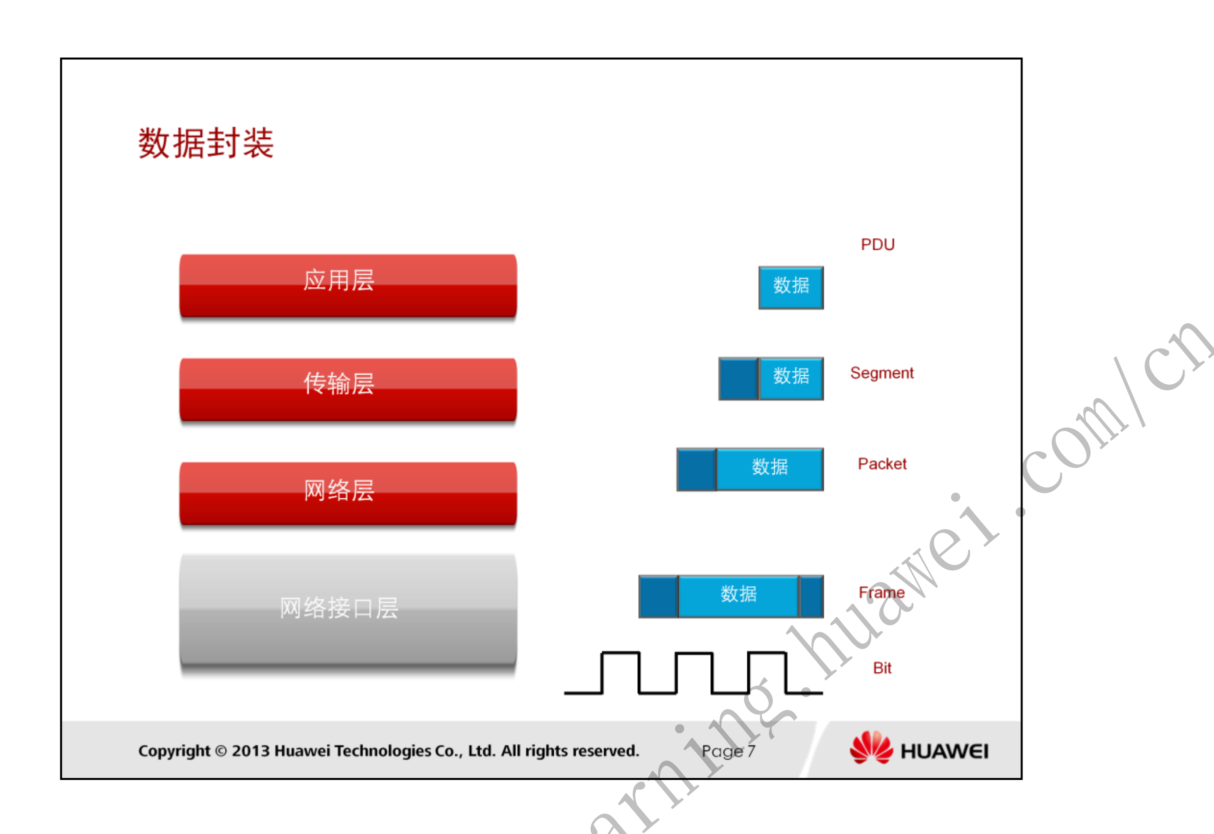
表示层：提供各种用于应用层数据的编码和转换功能，确保一个系统的应用层发送的数据能被另一个系统的应用层识别。

应用层：OSI参考模型中最靠近用户的一层，为应用程序提供网络服务。

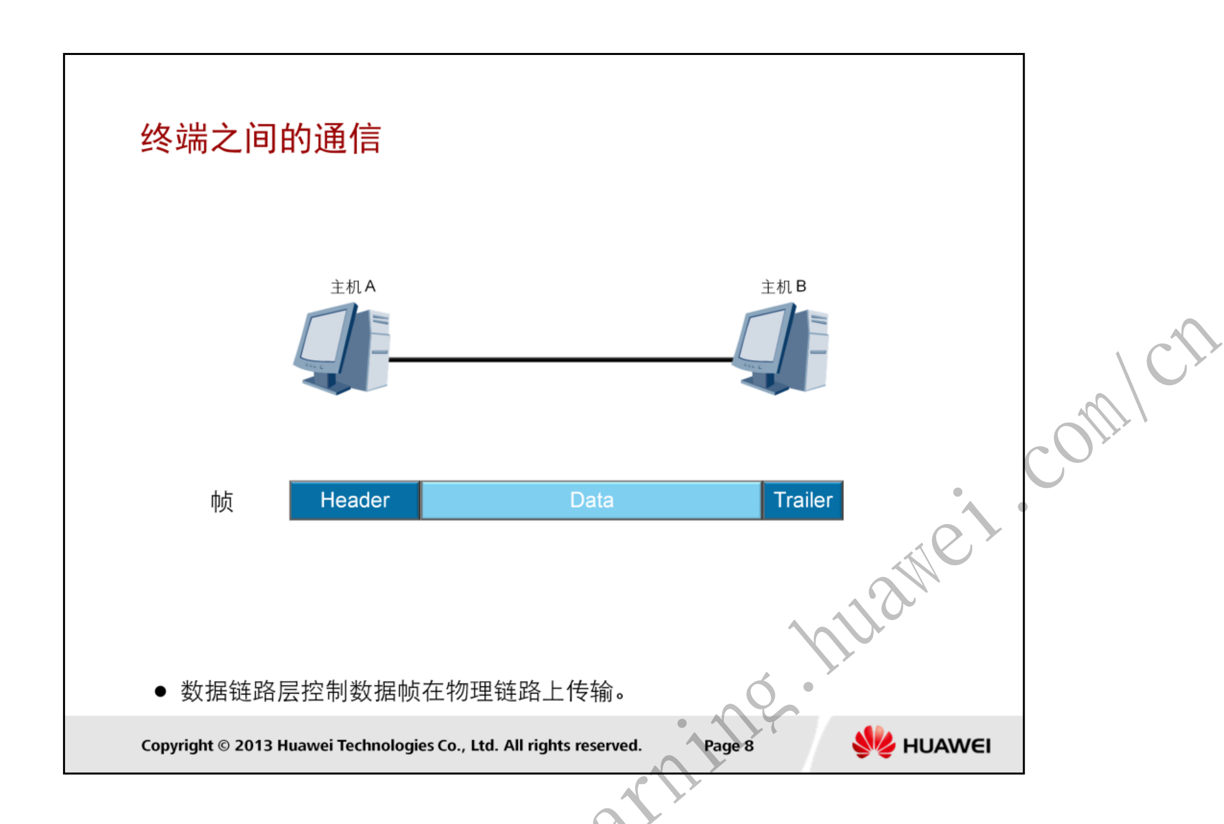


TCP/IP模型同样采用了分层结构，层与层相对独立但是相互之间也具备非常密切的协作关系。

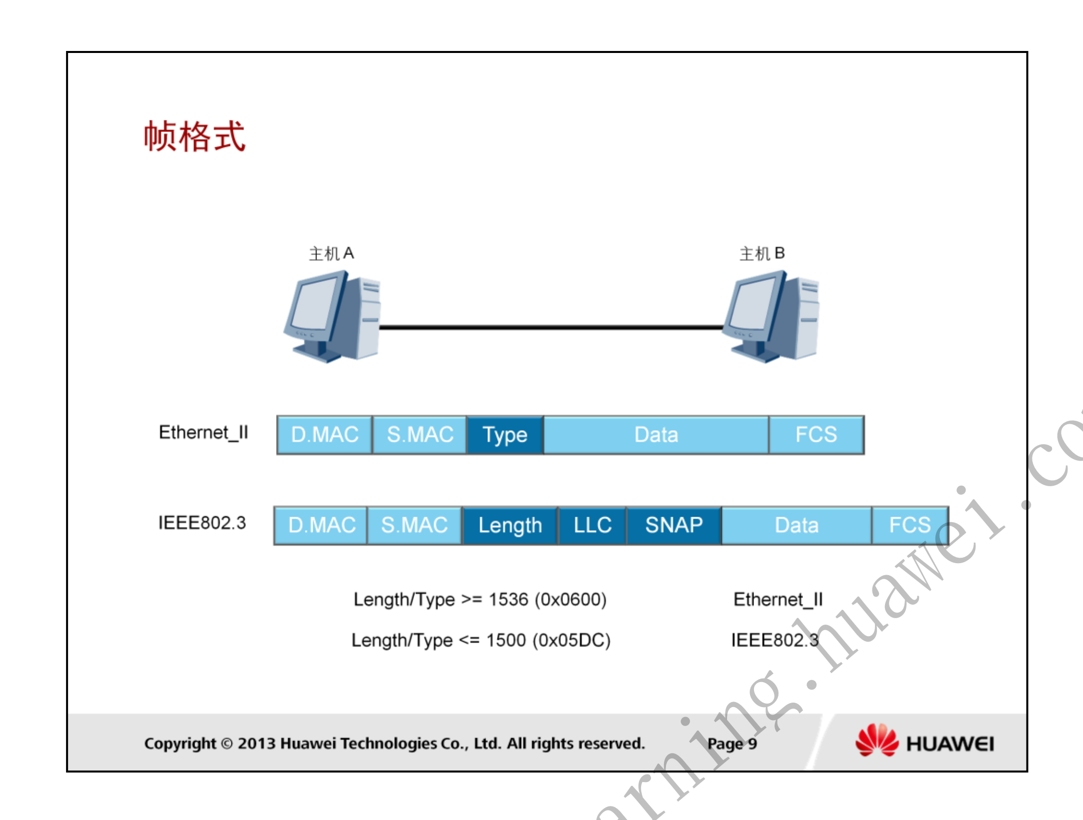
TCP/IP模型将网络分为四层。TCP/IP模型不关注底层物理介质，主要关注终端之间的逻辑数据流转发。TCP/IP模型的核心是网络层和传输层：网络层解决网络之间的逻辑转发问题，传输层保证源端到目的端之间的可靠传输。最上层的应用层通过各种协议向终端用户提供业务应用。



应用数据需要经过TCP/IP每一层处理之后才能通过网络传输到目的端，每一层上都使用该层的协议数据单元PDU（Protocol Data Unit）彼此交换信息。不同层的PDU中包含有不同的信息，因此PDU在不同层被赋予了不同的名称。如上层数据在传输层添加TCP报头后得到的PDU被称为Segment（数据段 ）；数据段被传递给网络层，网络层添加IP报头得到的PDU被称为Packet（数据包）；数据包被传递到数据链路层，封装数据链路层报头得到的PDU被称为Frame（数据帧）；最后，帧被转换为比特，通过网络介质传输。这种协议栈逐层向下传递数据，并添加报头和报尾的过程称为封装。



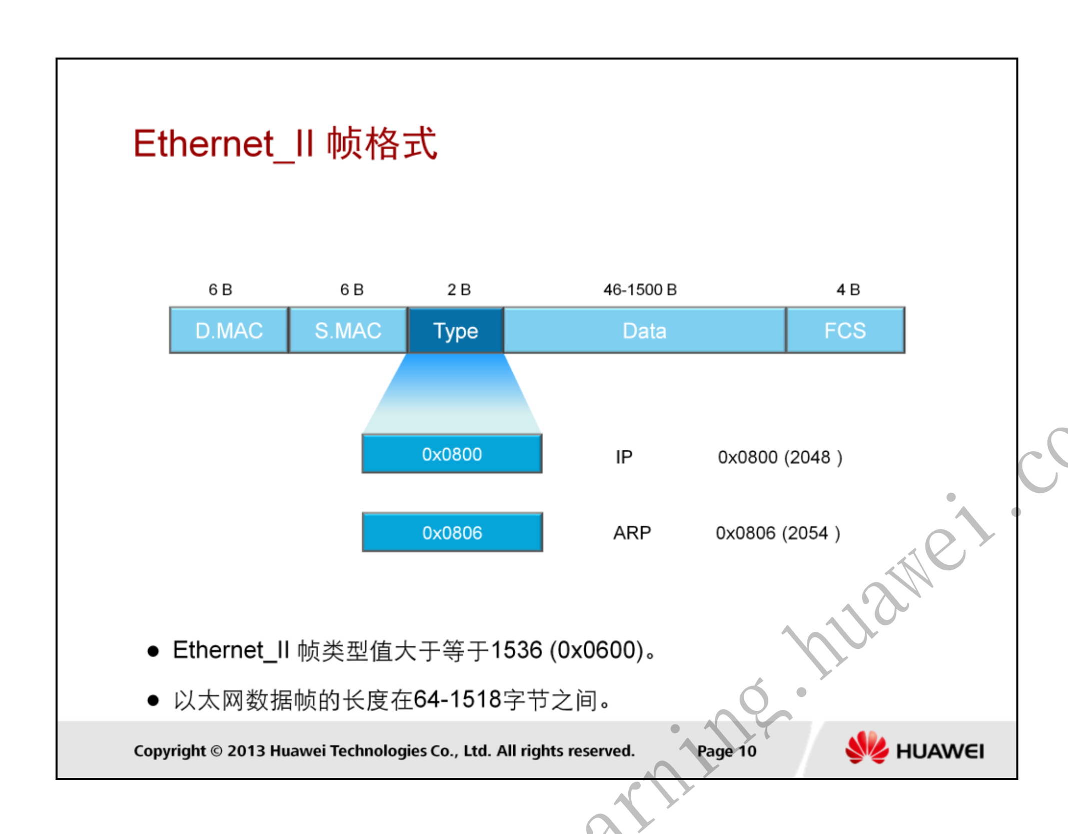
数据包在以太网物理介质上传播之前必须封装头部和尾部信息。封装后的数据包称为称为数据帧，数据帧中封装的信息决定了数据如何传输。以太网上传输的数据帧有两种格式，选择哪种格式由TCP/IP协议簇中的网络层决定。



以太网上使用两种标准帧格式。第一种是上世纪80年代初提出的DIX v2格式，即Ethernet II帧格式。Ethernet II后来被IEEE 802标准接纳，并写进了IEEE 802.3x-1997的3.2.6节。第二种是1983年提出的IEEE 802.3格式。这两种格式的主要区别在于，Ethernet II格式中包含一个Type字段，标识以太帧处理完成之后将被发送到哪个上层协议进行处理。IEEE 802.3格式中，同样的位置是长度字段。

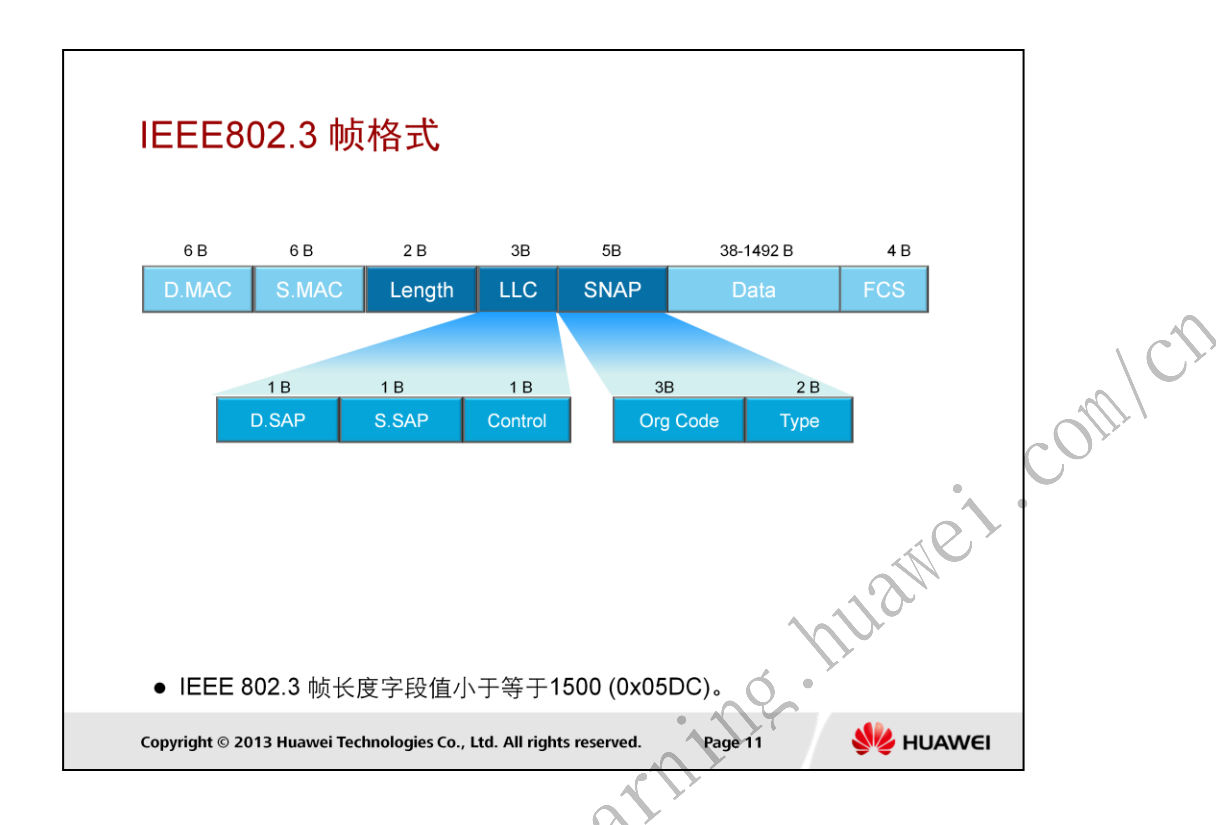
不同的Type字段值可以用来区别这两种帧的类型，当Type字段值小于等于1500（或者十六进制的0x05DC）时，帧使用的是IEEE 802.3格式。当Type字段值大于等于1536 （或者十六进制的0x0600）时，帧使用的是Ethernet II格式。以太网中大多数的数据帧使用的是Ethernet II格式。

以太帧中还包括源和目的MAC地址，分别代表发送者的MAC和接收者的MAC，此外还有帧校验序列字段，用于检验传输过程中帧的完整性。



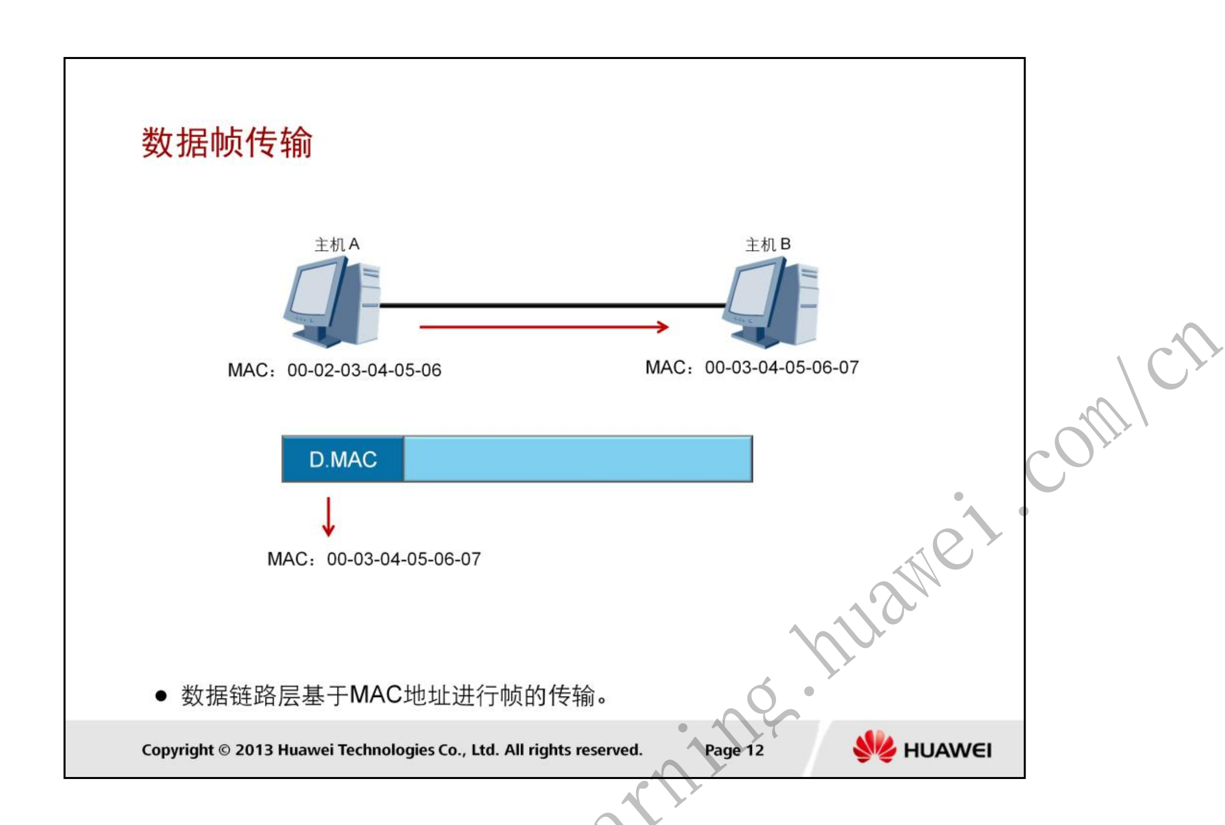
Ethernet_II的帧中各字段说明如下：

1. DMAC（Destination MAC）是目的MAC地址。DMAC字段长度为6个字节，标识帧的接收者。
2. SMAC（Source MAC）是源MAC地址。SMAC字段长度为6个字节，标识帧的发送者。
3. 类型字段（Type）用于标识数据字段中包含的高层协议，该字段长度为2个字节。类型字段取值为0x0800的帧代表IP协议帧；类型字段取值为0806的帧代表ARP协议帧。
4. 数据字段(Data)是网络层数据，最小长度必须为46字节以保证帧长至少为64字节，数据字段的最大长度为1500字节。
5. 循环冗余校验字段（FCS）提供了一种错误检测机制。该字段长度为4个字节。

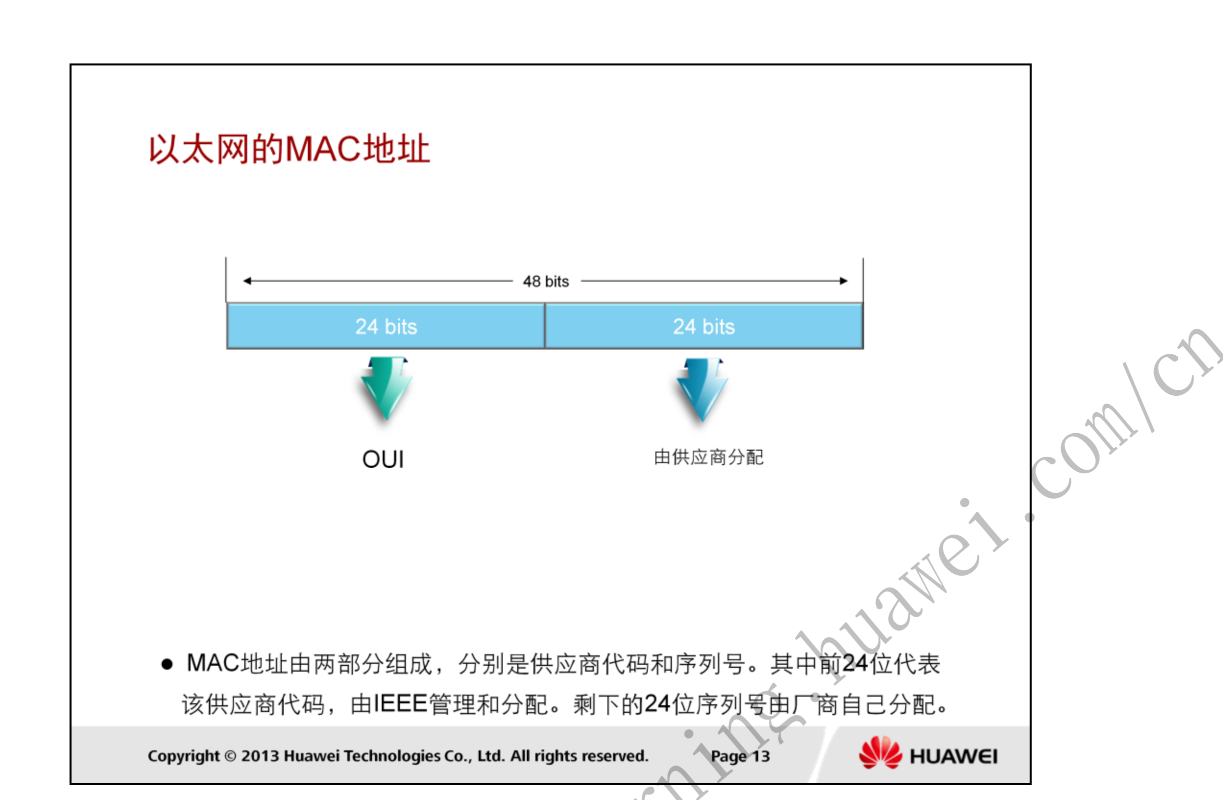


EEE802.3帧格式类似于Ethernet_II帧，只是Ethernet_II帧的Type域被802.3帧的Length域取代，并且占用了Data字段的8个字节作为LLC和SNAP字段。

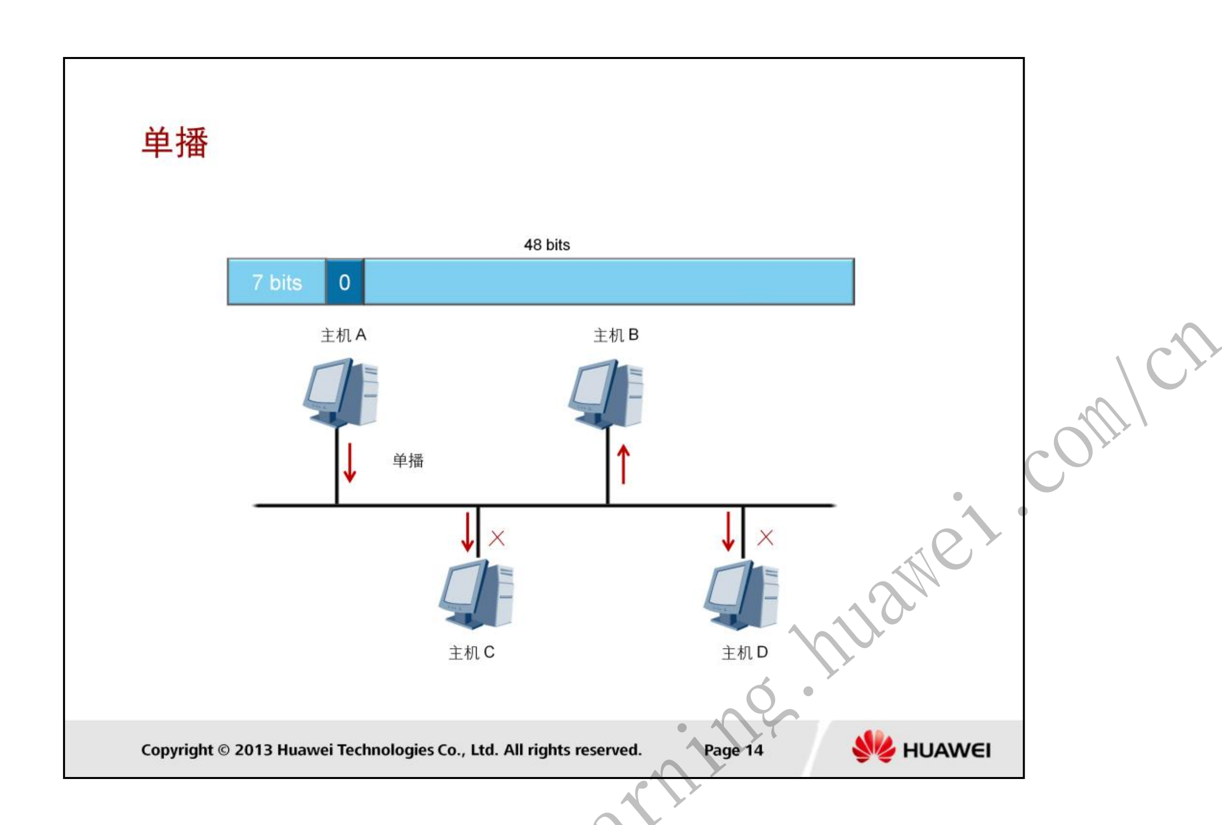
1. Length字段定义了Data字段包含的字节数。
2. 逻辑链路控制LLC（Logical Link Control）由目的服务访问点DSAP（Destination Service Access Point）、源服务访问点SSAP（Source Service Access Point）和Control字段组成。
3. SNAP（Sub-network Access Protocol）由机构代码（Org Code）和类型（Type）字段组成。Org code三个字节都为0。Type字段的含义与Ethernet_II帧中的Type字段相同。IEEE802.3帧根据DSAP和SSAP字段的取值又可分为以下几类:
   1) 当DSAP和SSAP都取特定值0xff时，802.3帧就变成了Netware-ETHERNET帧，用来承载NetWare类型的数据。
   2）当DSAP和SSAP都取特定值0xaa时，802.3帧就变成了ETHERNET_SNAP帧。ETHERNET_SNAP帧可以用于传输多种协议。
   3） DSAP和SSAP其他的取值均为纯IEEE802.3帧。



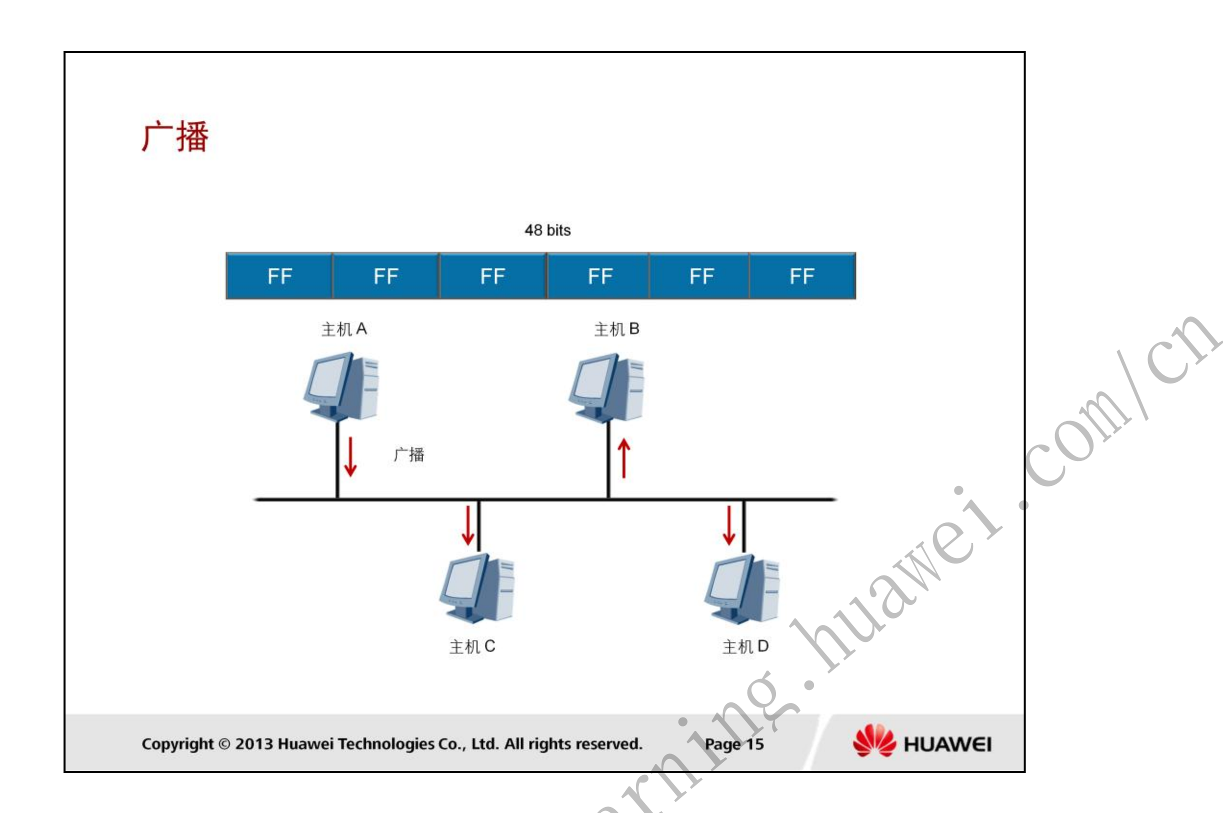
以太网在二层链路上通过MAC地址来唯一标识网络设备，并且实现局域网上网络设备之间的通信。MAC地址也叫物理地址，大多数网卡厂商把MAC地址烧入了网卡的ROM中。发送端使用接收端的MAC地址作为目的地址。以太帧封装完成后会通过物理层转换成比特流在物理介质上传输。



如同每一个人都有一个名字一样，每一台网络设备都用物理地址来标识自己，这个地址就是MAC地址。网络设备的MAC地址是全球唯一的。MAC地址长度为48比特，通常用十六进制表示。MAC地址包含两部分：前24比特是组织唯一标识符（OUI，Organizationally Unique Identifier），由IEEE统一分配给设备制造商。例如，华为的网络产品的MAC地址前24比特是0x00e0fc。后24位序列号是厂商分配给每个产品的唯一数值，由各个厂商自行分配（这里所说的产品可以是网卡或者其他需要MAC地址的设备）。



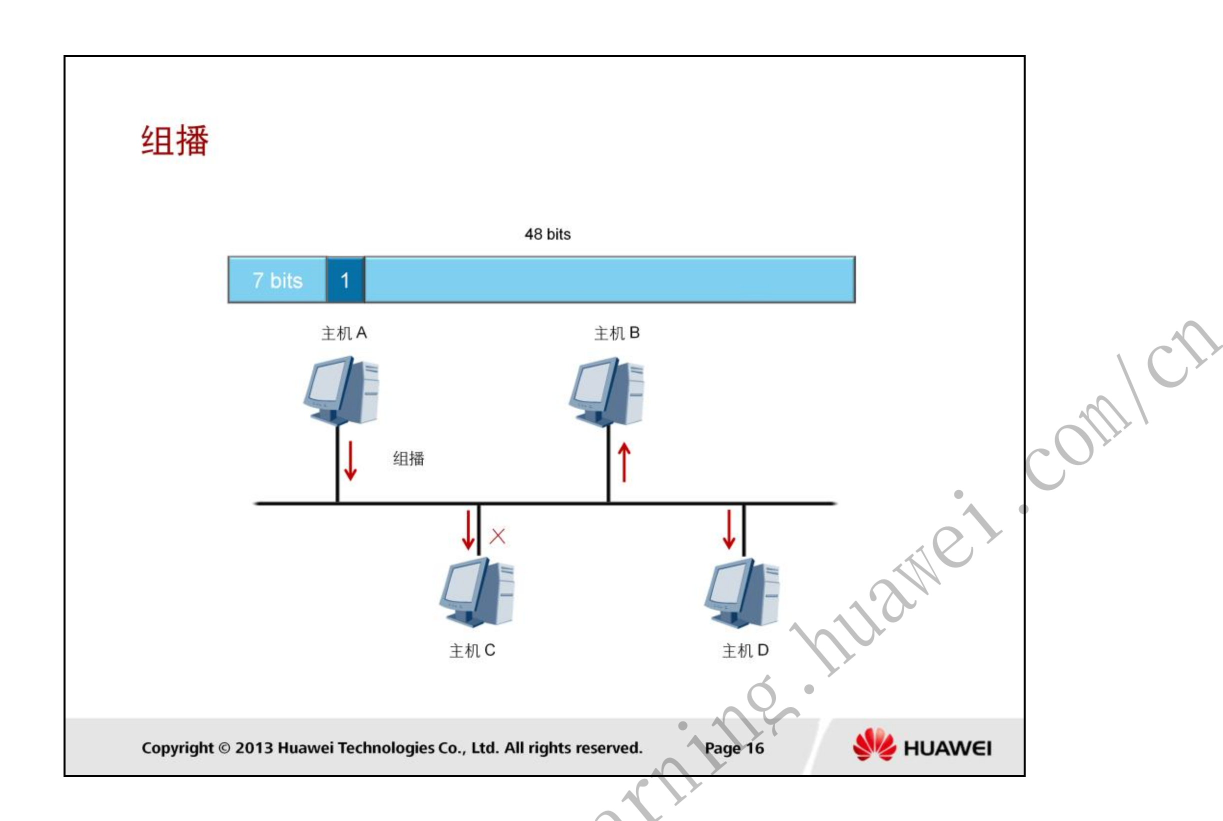
局域网上的帧可以通过三种方式发送。第一种是单播，指从单一的源端发送到单一的目的端。每个主机接口由一个MAC地址唯一标识，MAC地址的OUI中，第一字节第8个比特表示地址类型。对于主机MAC地址，这个比特固定为0，表示目的MAC地址为此MAC地址的帧都是发送到某个唯一的目的端。在冲突域中，所有主机都能收到源主机发送的单播帧，但是其他主机发现目的地址与本地MAC地址不一致后会丢弃收到的帧，只有真正的目的主机才会接收并处理收到的帧。



第二种发送方式是广播，表示帧从单一的源发送到共享以太网上的所有主机。广播帧的目的MAC地址为十六进制的FF:FF:FF:FF:FF:FF，所有收到该广播帧的主机都要接收并处理这个帧。

广播方式会产生大量流量，导致带宽利用率降低，进而影响整个网络的性能。

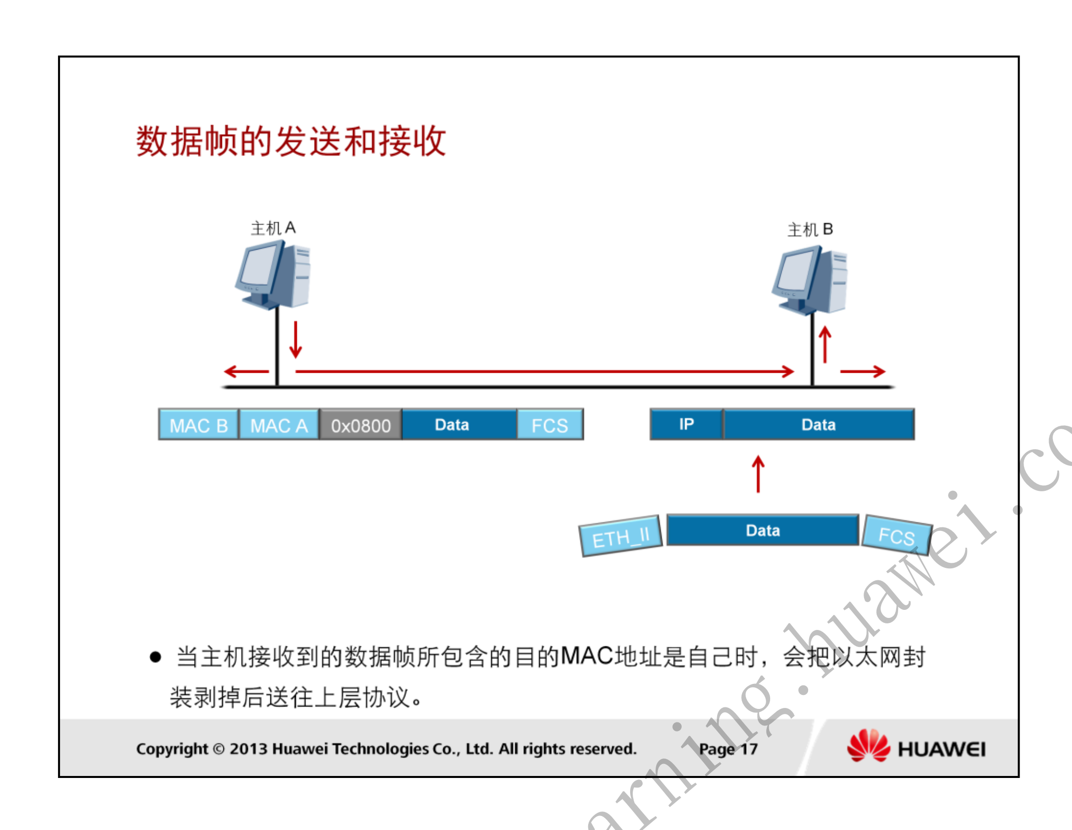
当需要网络中的所有主机都能接收到相同的信息并进行处理的情况下，通常会使用广播方式。



第三种发送方式为组播，组播比广播更加高效。组播转发可以理解为选择性的广播，主机监听特定组播地址，接收并处理目的MAC地址为该组播MAC地址的帧。

组播MAC地址和单播MAC地址是通过第一字节中的第8个比特区分的。组播MAC地址的第8个比特为1，而单播MAC地址的第8个比特为0。

当需要网络上的一组主机（而不是全部主机）接收相同信息，并且其他主机不受影响的情况下通常会使用组播方式。



帧从主机的物理接口发送出来后，通过传输介质传输到目的端。共享网络中，这个帧可能到达多个主机。主机检查帧头中的目的MAC地址，如果目的MAC地址不是本机MAC地址，也不是本机侦听的组播或广播MAC地址，则主机会丢弃收到的帧。

如果目的MAC地址是本机MAC地址，则接收该帧，检查帧校验序列（FCS）字段，并与本机计算的值对比来确定帧在传输过程中是否保持了完整性。如果帧的FCS值与本机计算的值不同，主机会认为帧已被破坏，并会丢弃该帧。如果该帧通过了FCS校验，则主机会根据帧头部中的Type字段来确定将帧发送给上层哪个协议处理。本例中，Type字段的值为0x0800，表明该帧需要发送到IP协议上处理。在发送给IP协议之前，帧的头部和尾部会被剥掉。



总结

- 网络设备如何确定以太网数据帧的上层协议？
- 终端设备接收到数据帧时，会如何处理？

 Page 18 

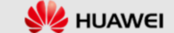
1. 以太网帧中包含一个Type字段，表示帧中的数据应该发送到上层哪个协议处理。比如，IP协议对应的Type值为0x0800，ARP协议对应的Type值为0x0806。
2. 主机检查帧头中的目的MAC地址，如果目的MAC地址不是本机MAC地址，也不是本机侦听的组播或广播MAC地址，则主机会丢弃收到的帧。如果目的MAC地址是本机MAC地址，则接收该帧，检查帧校验序列（FCS）字段，并与本机计算的值对比来确定帧在传输过程中是否保持了完整性。如果检查通过，就会剥离帧头和帧尾，然后根据帧头中的Type字段来决定把数据发送到哪个上层协议进行后续处理。


谢谢
www.huawei.com


IP编址
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



网络层位于数据链路层与传输层之间。网络层中包含了许多协议，其中最为重要的协议就是IP协议。网络层提供了IP路由功能。理解IP路由除了要熟悉IP协议的工作机制之外，还必须理解IP编址以及如何合理地使用IP地址来设计网络。

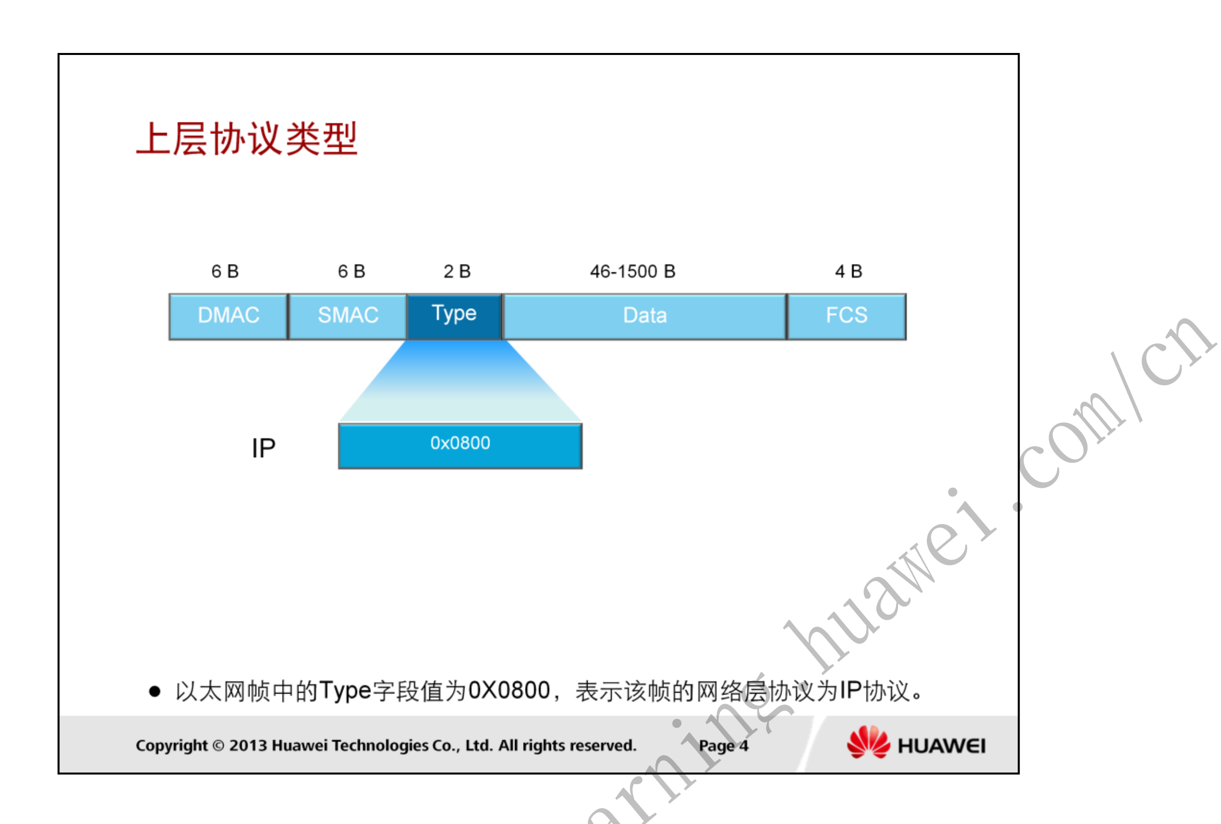




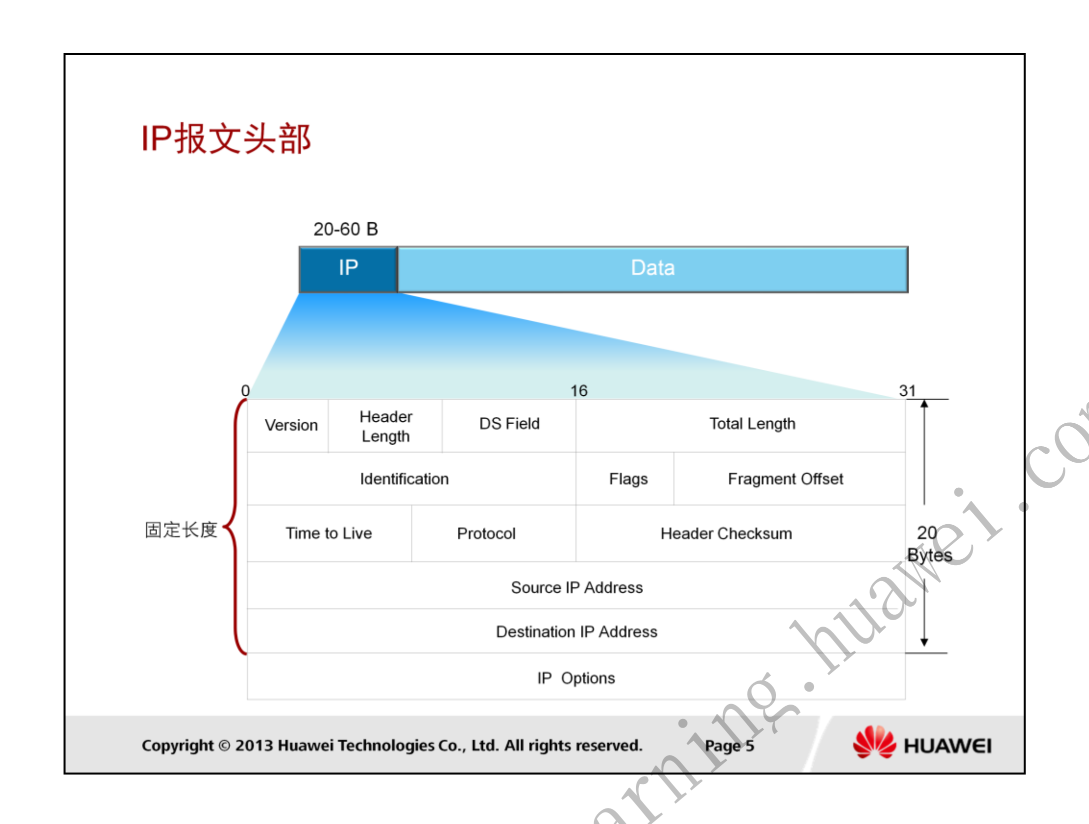
学习目标

学完本课程后，您应该能：

- 掌握IP报文的结构
- 掌握公有IP地址、私有IP地址以及特殊IP地址的范围
- 掌握VLSM技术
- 理解网关的作用



在剥掉帧的头部和尾部之前，网络设备需要根据帧头中Type字段确定下一步将帧发送到哪个上层协议进行处理。本例中的帧头部Type字段表示该帧需要上送到IP协议进行处理。以下将介绍帧的头部和尾部被剥掉后，IP协议将如何处理帧中的数据。



IP报文头部信息用于指导网络设备对报文进行路由和分片。同一个网段内的数据转发通过链路层即可实现，而跨网段的数据转发需要使用网络设备的路由功能。分片是指数据包超过一定长度时，需要被划分成不同的片段使其能够在网络中传输。

IP报文头部长度为20到60字节，报文头中的信息可以用来指导网络设备如何将报文从源设备发送到目的设备。其中，版本字段表示当前支持的IP协议版本，当前的版本号为4。DS字段早期用来表示业务类型，现在用于支持QoS中的差别服务模型，实现网络流量优化。

源和目的IP地址是分配给主机的逻辑地址，用于在网络层标识报文的发送方和接收方。根据源和目的IP地址可以判断目的端是否与发送端位于同一网段，如果二者不在同一网段，则需要采用路由机制进行跨网段转发。

IP编址

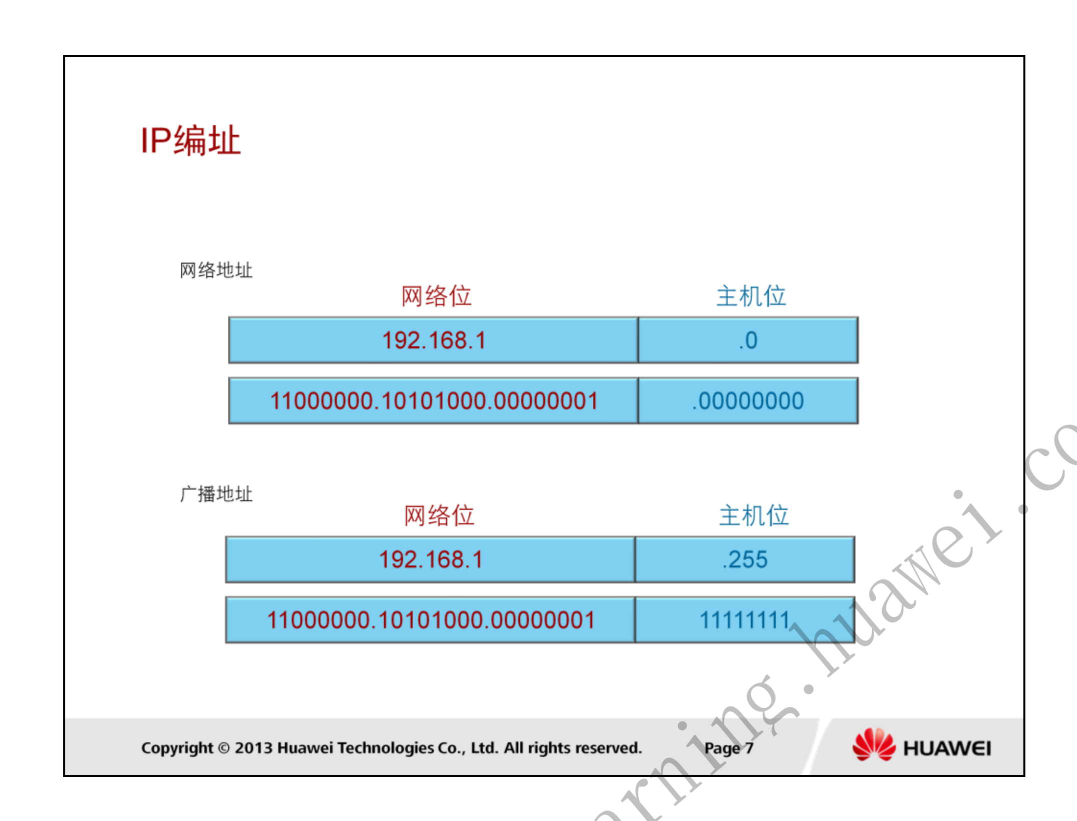
网络位	主机位
192.168.1	.1
11000000.10101000.00000001	.00000001

- IP地址分为网络部分和主机部分。
- IP地址由32个二进制位组成，通常用点分十进制形式表示。

 Page 6



IPv4地址为32比特的二进制数，通常用点分十进制表示。IP地址用来标识网络中的设备，具有IP地址的设备可以在同一网段内或跨网段通信。IP地址包括两部分，第一部分是网络号，表示IP地址所属的网段，第二部分是主机号，用来唯一标识本网段上的某台网络设备。



每个网段上都有两个特殊地址不能分配给主机或网络设备。第一个是该网段的网络地址，该IP地址的主机位为全0，表示一个网段。第二个地址是该网段中的广播地址，目的地址为广播地址的报文会被该网段中的所有网络设备接收。广播地址的主机位为全1。除网络地址和广播地址以外的其他IP地址可以作为网络设备的IP地址。

二进制、十进制和十六进制

进制	字符范围	基值
二进制	0—1	2
十进制	0—9	10
十六进制	0—9，A—F	16

- 在IP网络中，二进制和十六进制是常用的编码方式。

 Page 8 

网络中的数据可以采用二进制、十进制、或十六进制来表示，了解这些进制对理解IP网络基础知识很有必要。每种进制使用不同的基值表示每一位的数值。二进制每一位只有0和1两个值，基值为2，二进制数的每一位都可以用2的x次幂来表示，x表示二进制数的位数。十六进制的每一位可以有16个数值，范围为0-F（即0-9和A-F），A对应十进制的10，F对应十进制的15（二进制的1111）。

进制之间转换

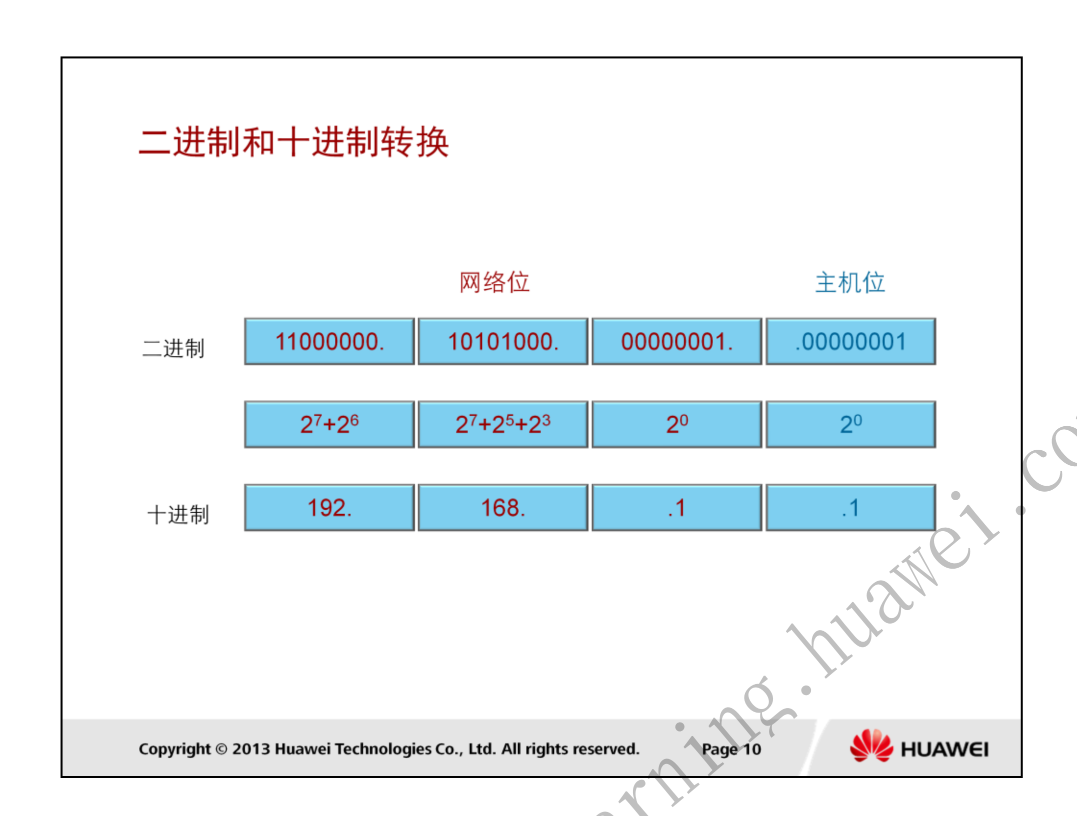
比特位	1	1	1	1	1	1	1	1
乘方	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
数值	128	64	32	16	8	4	2	1

十进制	二进制	十六进制
0	00000000	00
1	00000001	01
2	00000010	02
3	00000011	03
4	00000100	04
5	00000101	05
6	00000110	06
7	00000111	07
8	00001000	08

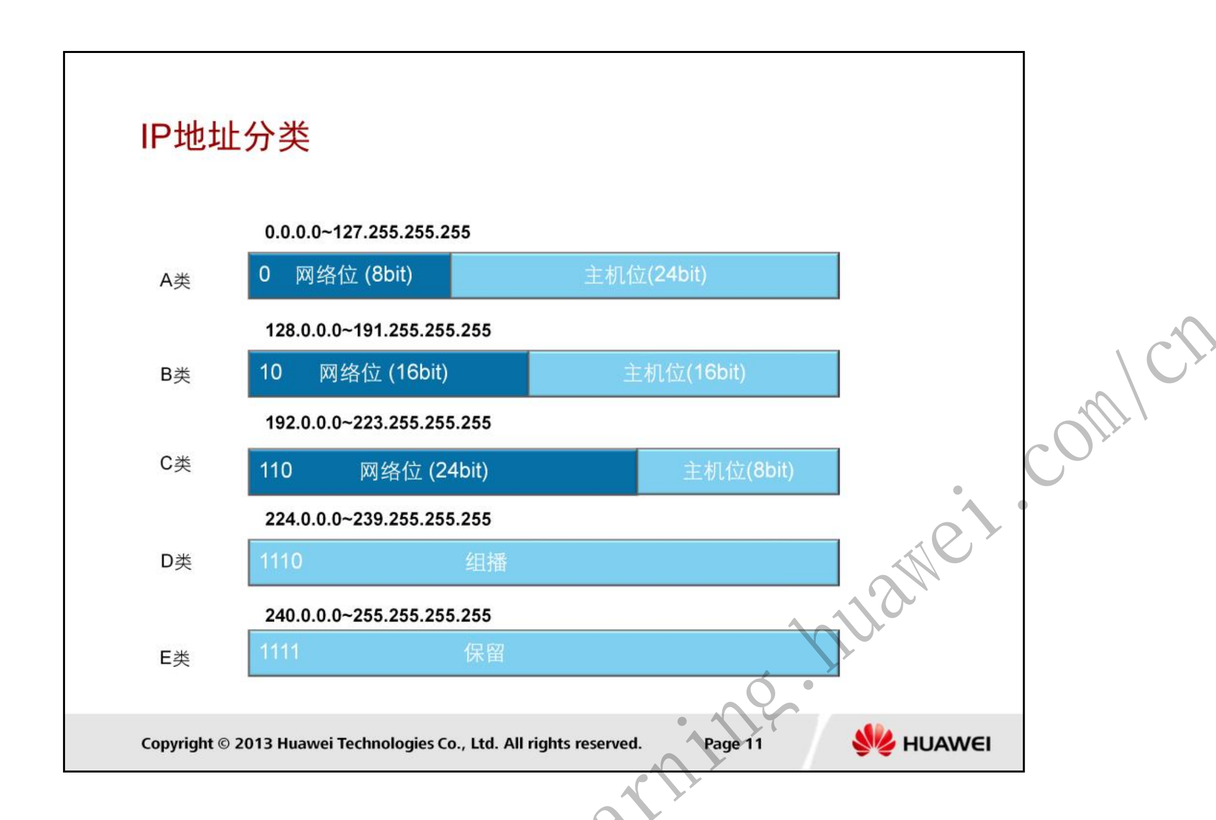
十进制	二进制	十六进制
9	00001001	09
10	00001010	0A
11	00001011	0B
12	00001100	0C
13	00001101	0D
14	00001110	0E
15	00001111	0F
…	…	…
255	11111111	FF



IP地址以字节为单位分为四段，每字节包含8个比特，可以表示0到255，共256个数值。从二进制到十进制转换表中可以看到每一位二进制数所代表的的十进制数。上面的表格举例说明了8位二进制数转换为十进制数和十六进制数的情况。从表格中也可以看到全0和全1所对应的十进制数和十六进制数。



32位的IP地址分为4个字节，每个字节有256个取值。因此，理论上IPv4可以有4,294,967,296个IP地址，但实际上只有其中一部分地址可以分配给网络设备使用。本例中，IP地址的前三个字节表示网络号，最后一个字节表示该网络上网络设备可用的地址范围。将二进制格式的IP地址转换为十进制格式时，需要把二进制中每一位1所代表的值加在一起，得出IP地址的十进制值。



IPv4地址被划分为A、B、C、D、E五类，每类地址的网络号包含不同的字节数。A类，B类，和C类地址为可分配IP地址，每类地址支持的网络数和主机数不同。比如，A类地址可支持126个网络，每个网络支持2^{24} （16,777,216)个主机地址，另外每个网段中的网络地址和广播地址不能分配给主机。C类地址支持200多万个网络，每个网络支持256个主机地址，其中254个地址可以分配给主机使用。

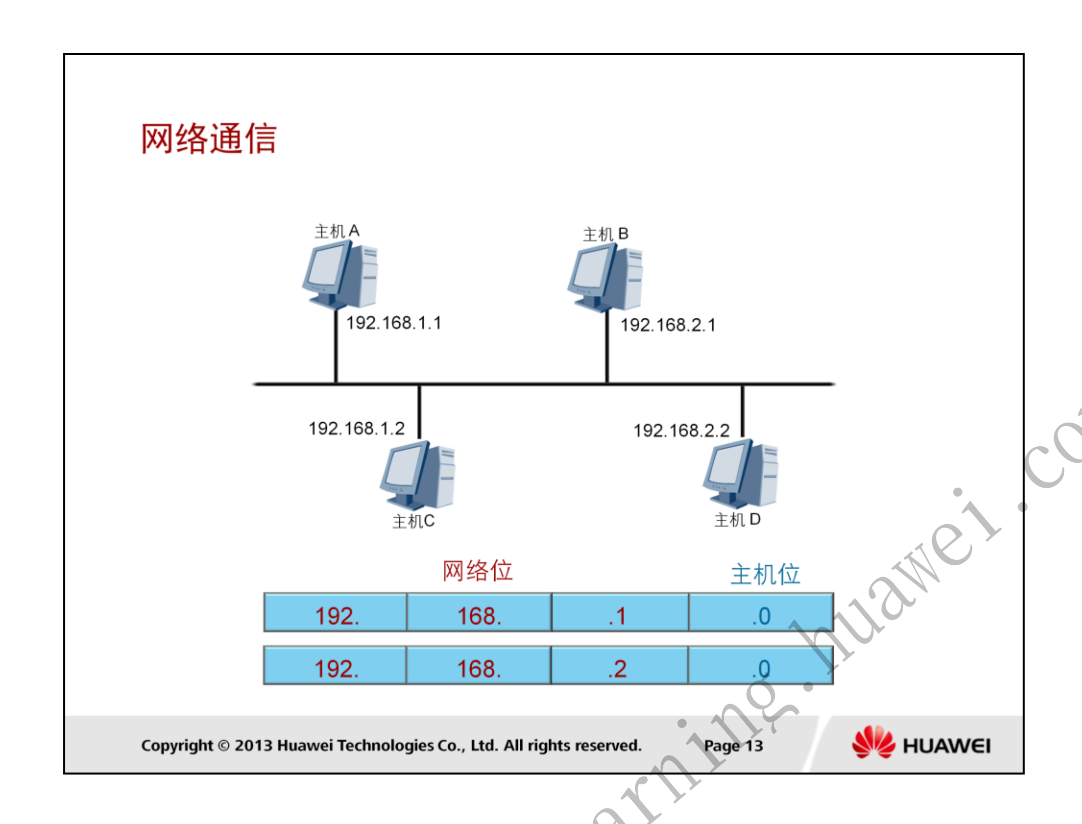
D类地址为组播地址。主机收到以D类地址为目的地址的报文后，且该主机是该组播组成员，就会接收并处理该报文。各类IP地址可以通过第一个字节中的比特位进行区分。如A类地址第一字节的最高位固定为0，B类地址第一字节的高两位固定为10，C类地址第一字节的高三位固定为110，D类地址第一字节的高四位固定为1110，E类地址第一字节的高四位固定为1111。

IP地址类型

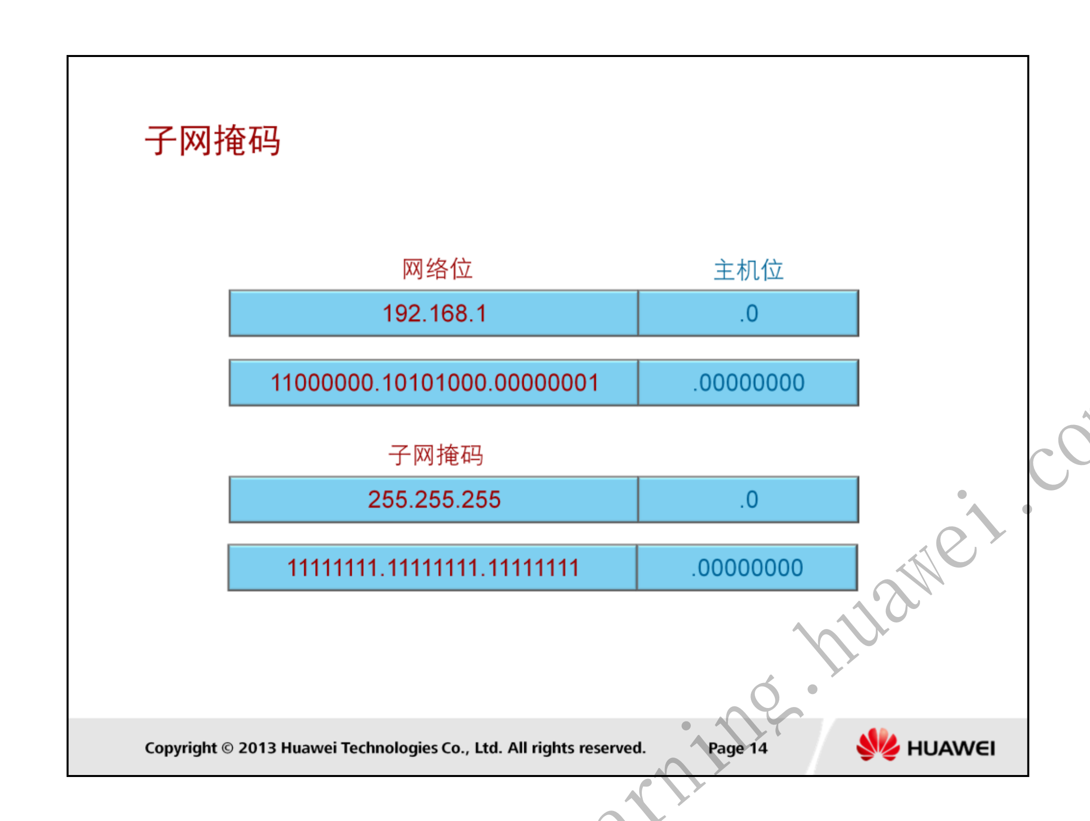
- 私有地址范围
  - 10.0.0.0~10.255.255.255
  - 172.16.0.0~172.31.255.255
  - 192.168.0.0~192.168.255.255
- 特殊地址
  - 127.0.0.0 ~ 127.255.255.255
  - 0.0.0.0
  - 255.255.255.255

 Page 12 

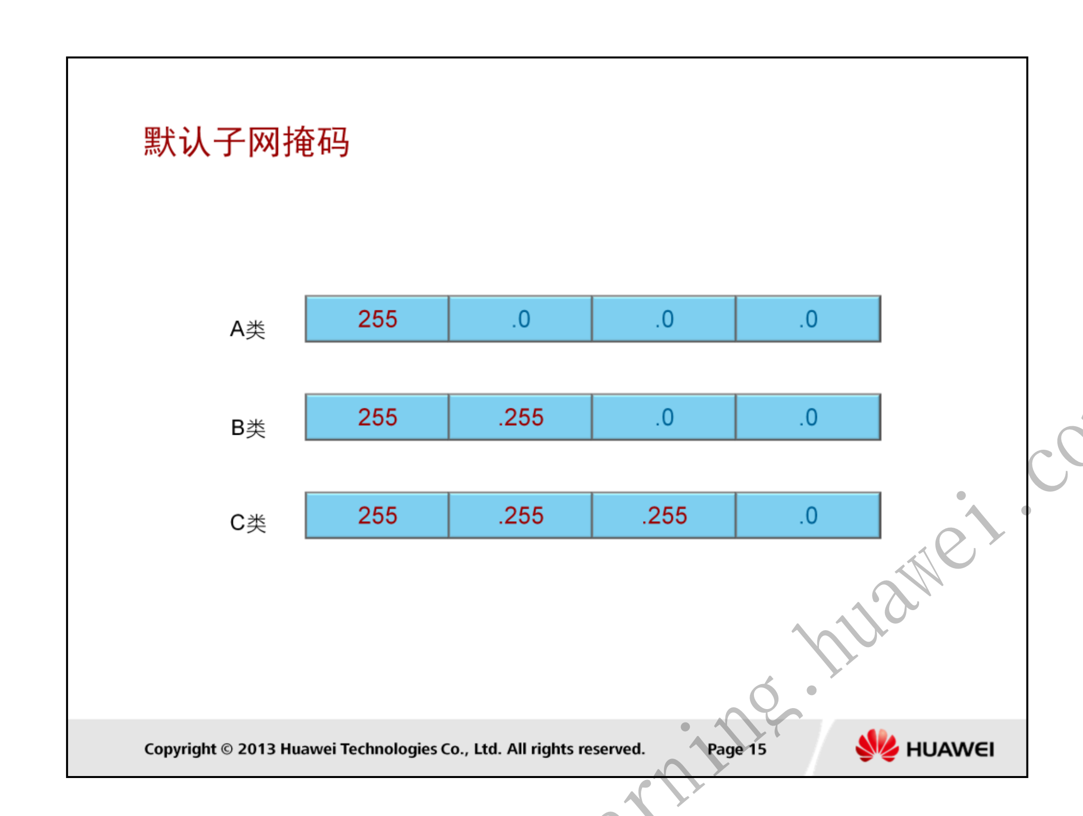
IPv4中的部分IP地址被保留用作特殊用途。为节省IPv4地址，A， B，C类地址段中都预留了特定范围的地址作为私网地址。现在，世界上所有终端系统和网络设备需要的IP地址总数已经超过了32位IPv4地址所能支持的最大地址数4，294，967，296。为主机分配私网地址节省了公网地址，可以用来缓解IP地址短缺的问题。企业网络中普遍使用私网地址，不同企业网络中的私网地址可以重叠。默认情况下，网络中的主机无法使用私网地址与公网通信；当需要与公网通信时，私网地址必须转换成公网地址。还有其他一些特殊IP地址，如127.0.0.0网段中的地址为环回地址，用于诊断网络是否正常。IPv4中的第一个地址0.0.0.0表示任何网络，这个地址的作用将在路由原理中详细介绍。IPv4中的最后一个地址255.255.255.255是0.0.0.0网络中的广播地址。



源主机必须要知道目的主机的IP地址后才能将数据发送到目的地。源主机向其他目的主机发送报文之前，需要检查目的IP地址和源IP地址是否属于同一个网段。如果是，则报文将被下发到底层协议进行以太网封装处理。如果目的地址和源地址属于不同网段，则主机需要获取下一跳路由器的IP地址，然后将报文下发到底层协议处理。



子网掩码用于区分网络部分和主机部分。子网掩码与IP地址的表示方法相同。每个IP地址和子网掩码一起可以用来唯一的标识一个网段中的某台网络设备。子网掩码中的1表示网络位，0表示主机位。



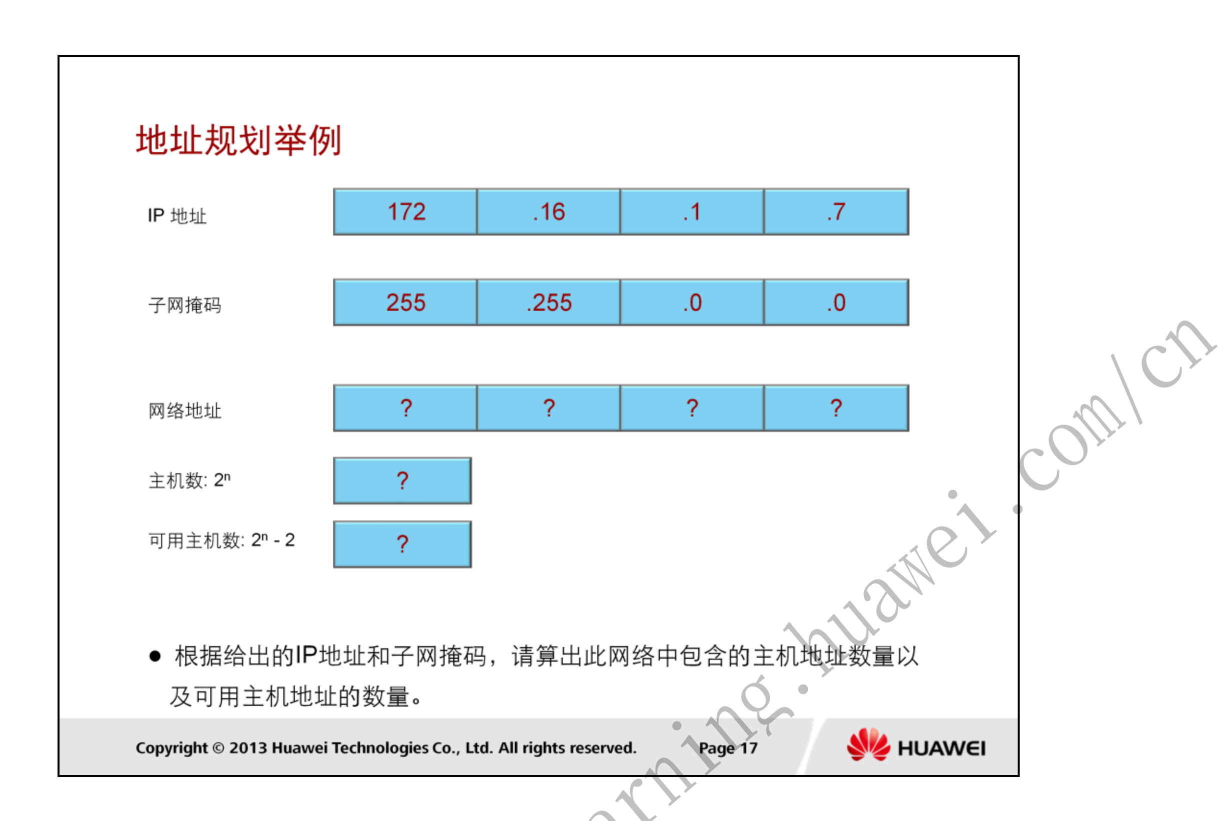
每类IP地址有一个缺省子网掩码。A类地址的缺省子网掩码为8位，即第一个字节表示网络位，其他三个字节表示主机位。B类地址的缺省子网掩码为16位，因此B类地址支持更多的网络，但是主机数也相应减少。C类地址的缺省子网掩码为24位，支持的网络最多，同时也限制了单个网络中主机的数量。

地址规划

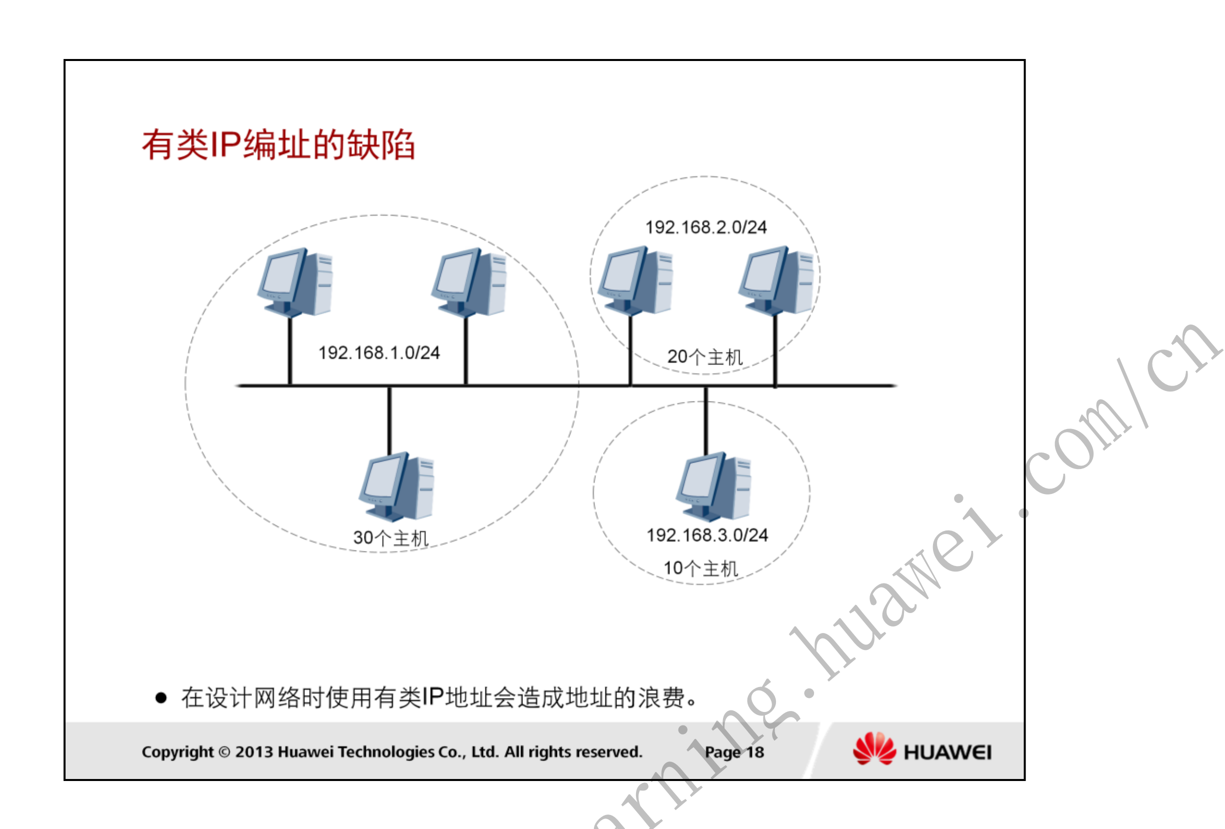
IP 地址	192	.168	.1	.7
子网掩码	255	.255	.255	.0
	11000000	10101000	00000001	00000111
	11111111	11111111	11111111	00000000
网络地址（二进制）	11000000	10101000	00000001	00000000
网络地址	192	.168	.1	.0
主机数: 2^n	256			
可用主机数: 2^n - 2	254			

 Page 16 HUAWEI

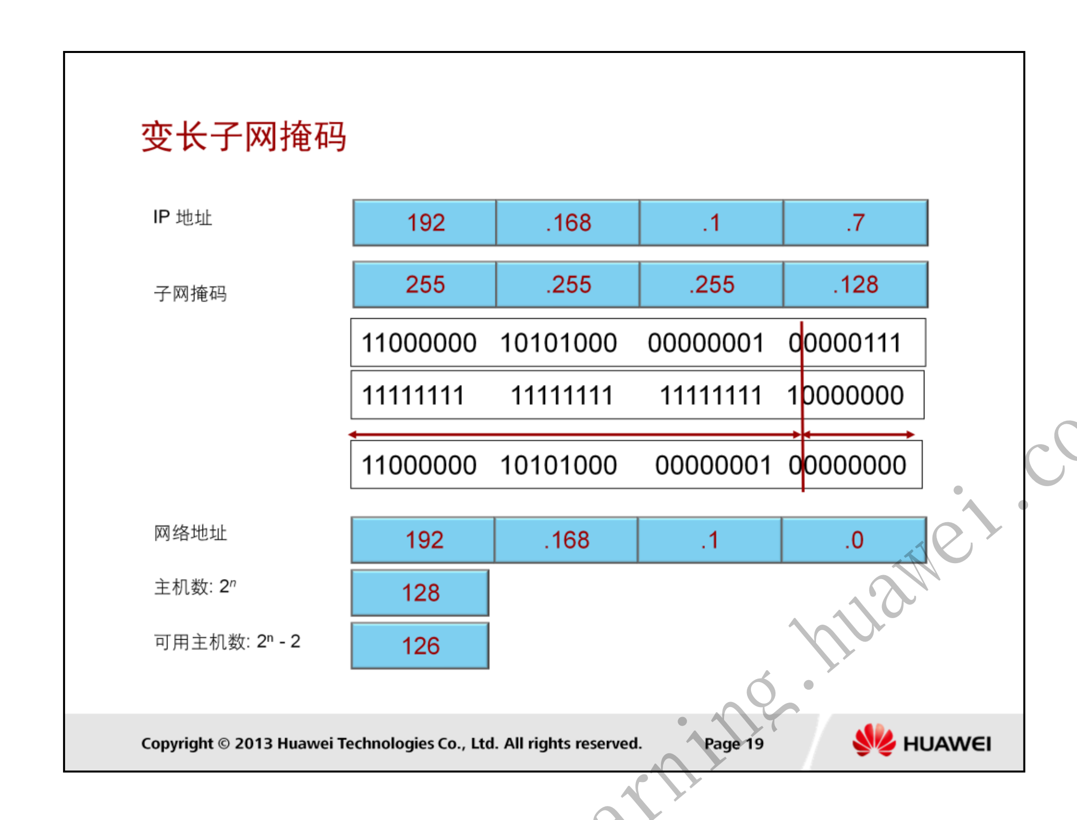
通过子网掩码可以判断主机所属的网段、网段上的广播地址、以及网段上支持的主机数。图中这个例子，主机地址为192.168.1.7，子网掩码为24位（C类IP地址的缺省掩码），从中我们可以判断该主机位于192.168.1.0/24网段。将IP地址中的主机位全部置为1，并转换为十进制数，即可得到该网段的广播地址192.168.1.255。网段中支持的主机数为2^n，n为主机位的个数。本例中n=8，2^8=256，减去本网段的网络地址和广播地址，可知该网段支持254个有效主机地址。



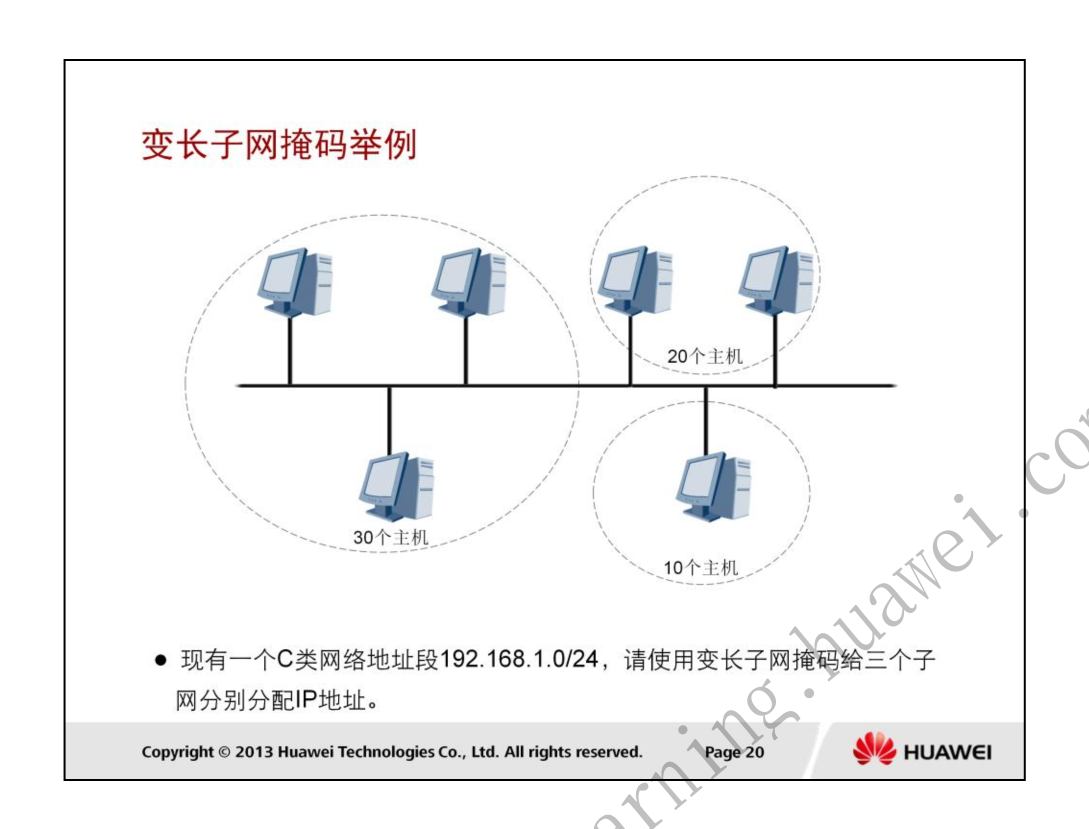
本例说明如何根据B类IP地址及其子网掩码判断主机所属的网段、网段中的广播地址、以及有效主机地址数量。判断过程与C类地址类似。



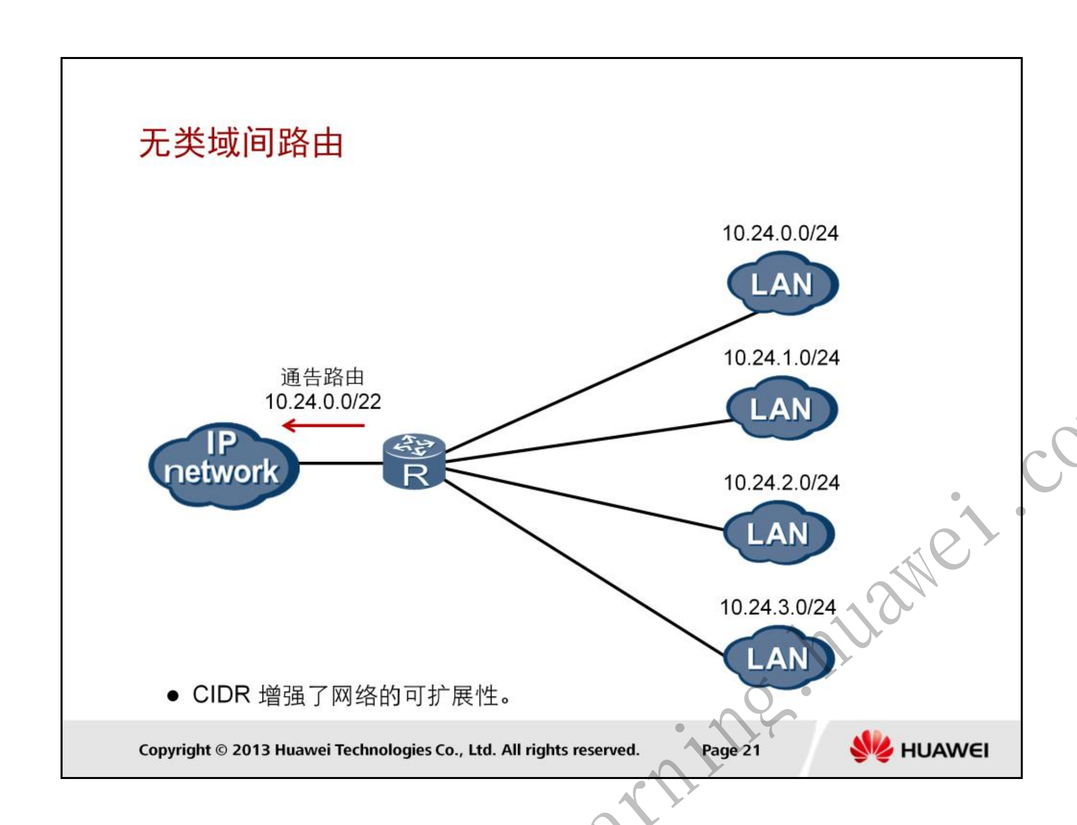
如果企业网络中希望通过规划多个网段来隔离物理网络上的主机，使用缺省子网掩码就会存在一定的局限性。网络中划分多个网段后，每个网段中的实际主机数量可能很有限，导致很多地址未被使用。如图所示的场景下，如果使用缺省子网掩码的编址方案，则地址使用率很低。



采用可变长子网掩码可解决上述问题。缺省子网掩码可以进一步划分，成为变长子网掩码（VLSM）。通过改变子网掩码，可以将网络划分为多个子网。本例中的地址为C类地址，缺省子网掩码为24位。现借用一个主机位作为网络位，借用的主机位变成子网位。一个子网位有两个取值0和1，因此可划分两个子网。该比特位设置为0，则子网号为0，该比特位设置为1，则子网号为128。将剩余的主机位都设置为0，即可得到划分后的子网地址；将剩余的主机位都设置为1，即可得到子网的广播地址。每个子网中支持的主机数为2^7-2（减去子网地址和广播地址），即126个主机地址。

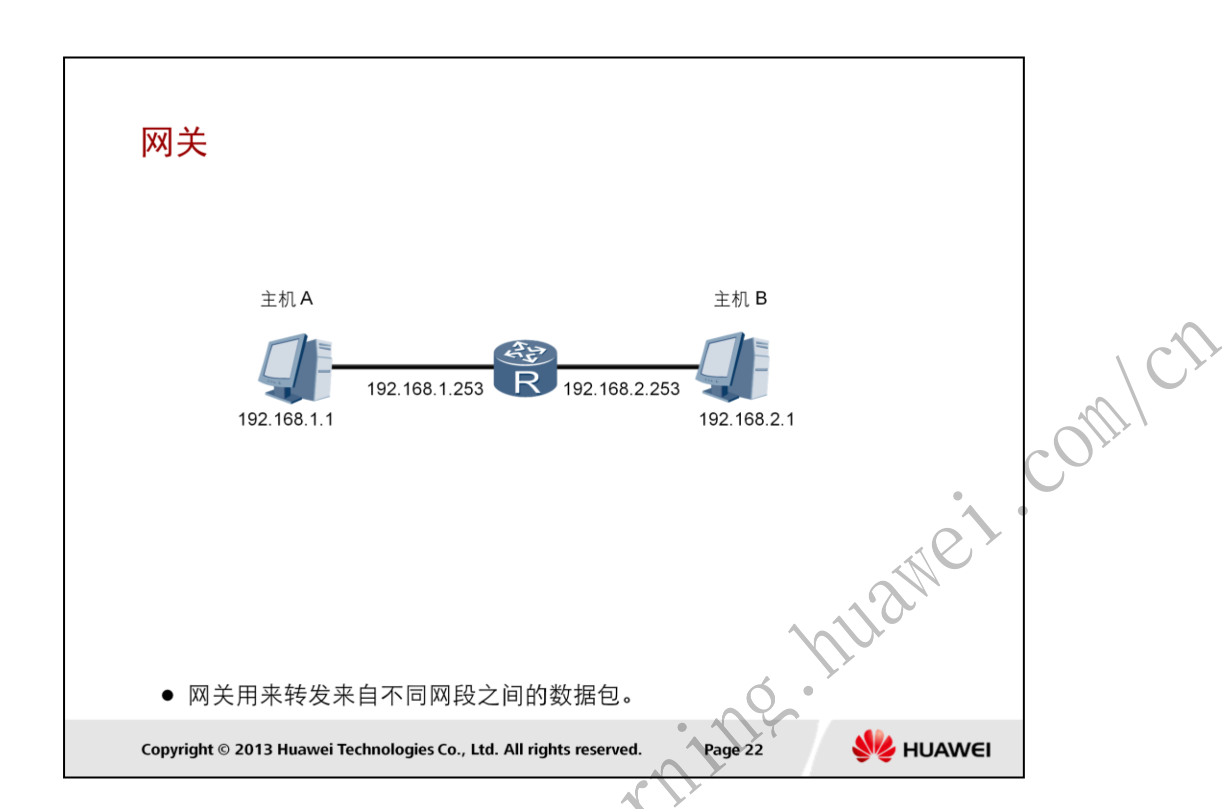


可变长子网掩码缓解了使用缺省子网掩码导致的地址浪费问题，同时也为企业网络提供了更为有效的编址方案。本例中需要使用可变长子网掩码来划分多个子网，借用一定数量的主机位作为子网位的同时，剩余的主机位必须保证有足够的IP地址供每个子网上的所有主机使用。



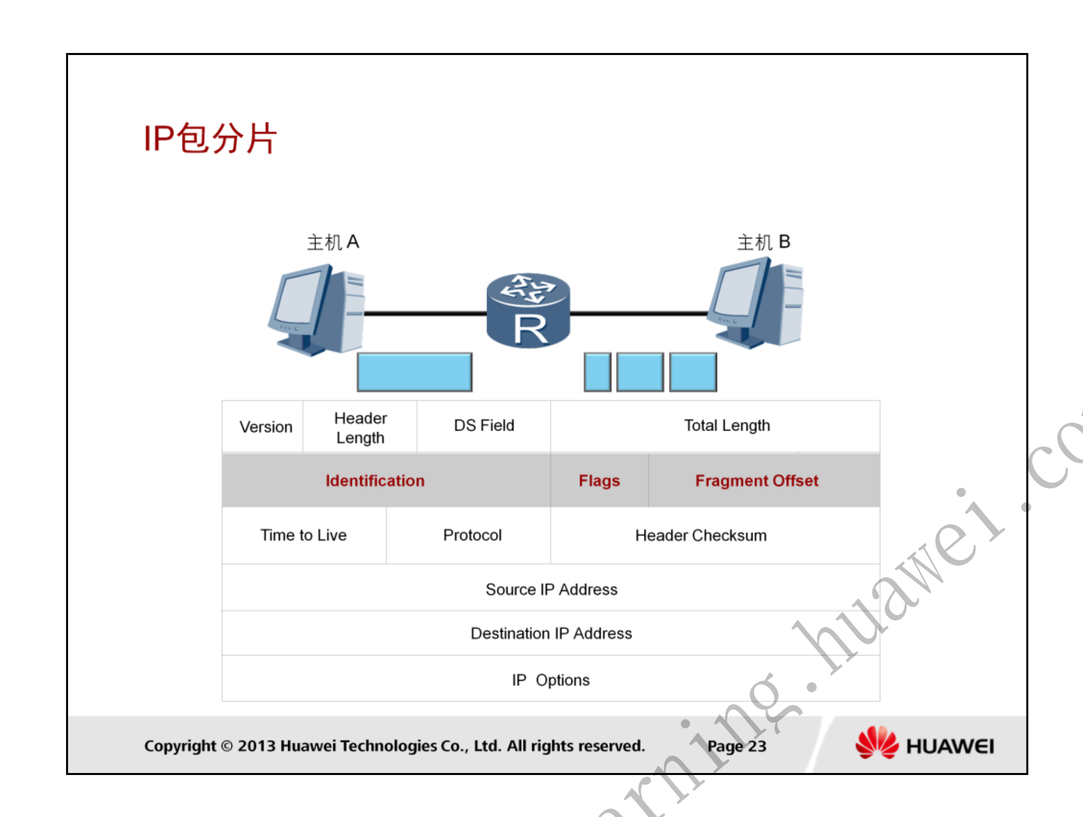
无类域间路由CIDR（Classless Inter Domain Routing）由RFC1817定义。CIDR突破了传统IP地址的分类边界，将路由表中的若干条路由汇聚为一条路由，减少了路由表的规模，提高了路由器的可扩展性。

如上图所示，一个企业分配到了一段A类网络地址，10.24.0.0/22。该企业准备把这些A类网络分配给各个用户群，目前已经分配了四个网段给用户。如果没有实施CIDR技术，企业路由器的路由表中会有四条下连网段的路由条目，并且会把它通告给其他路由器。通过实施CIDR技术，我们可以在企业的路由器上把这四条路由10.24.0.0/24，10.24.1.0/24，10.24.2.0/24，10.24.3.0/24汇聚成一条路由10.24.0.0/22。这样，企业路由器只需通告10.24.0.0/22这一条路由，大大减少了路由表的规模。



报文转发过程中，首先需要确定转发路径以及通往目的网段的接口，然后将报文封装在以太帧中通过指定的物理接口转发出去。如果目的主机与源主机不在同一网段，报文需要先转发到网关，然后通过网关将报文转发到目的网段。

网关是指接收并处理本地网段主机发送的报文并转发到目的网段的设备。为实现此功能，网关必须知道目的网段的IP地址。网关设备上连接本地网段的接口地址即为该网段的网关地址。



网络中转发的IP报文的长度可以不同，但如果报文长度超过了数据链路所支持的最大长度，则报文需要分割成若干个较小的片段才能够在链路上传输。将报文分割成多个片段的过程叫做分片。

接收端根据分片报文中的标识符（Identification），标志（Flags），及片偏移（Fragment Offset）字段对分片报文进行重组。标识符用于识别属于同一个数据包的分片，以区别于同一主机或其他主机发送的其它数据包分片，保证分片被正确的重新组合。标志字段用于判断是否已经收到最后一个分片。最后一个分片的标志字段设置为0，其他分片的标志字段设置为1，目的端在收到标志字段为0的分片后，开始重组报文。片偏移字段表示每个分片在原始报文中的位置。第一个分片的片偏移为0，第二个分片的片偏移表示紧跟第一个分片后的第一个比特的位置。比如，如果首片报文包含1259比特，那么第二片报文的片偏移字段值就应该为1260。

生存时间

主机 A　　　　R　　　　R　　　　主机 B

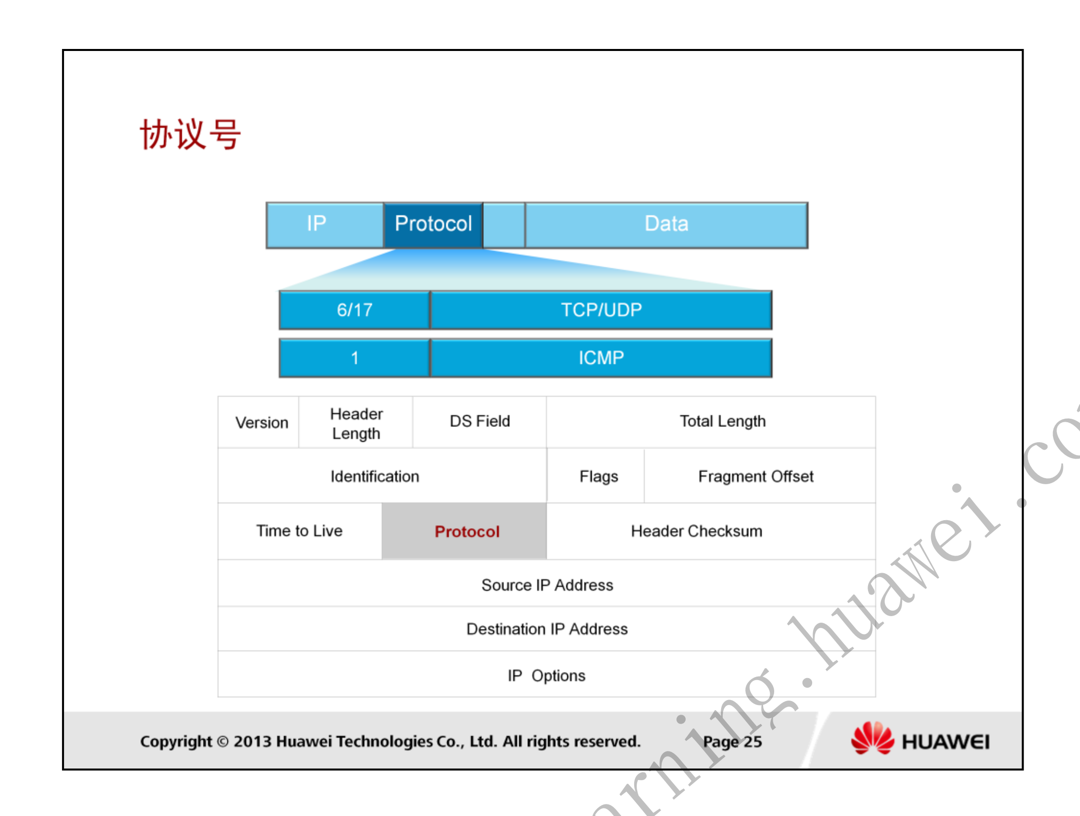
TTL=255　　TTL=254　　TTL=253

Version	Header Length	DS Field	Total Length	
Identification			Flags	Fragment Offset
Time to Live	Protocol	Header Checksum		
Source IP Address				
Destination IP Address				
IP Options				

 Page 24 HUAWEI

报文在网段间转发时，如果网络设备上的路由规划不合理，就可能会出现环路，导致报文在网络中无限循环，无法到达目的端。环路发生后，所有发往这个目的地的报文都会被循环转发，随着这种报文逐渐增多，网络将会发生拥塞。

为避免环路导致的网络拥塞，IP报文头中包含一个生存时间TTL（Time To Live）字段。报文每经过一台三层设备，TTL值减1。初始TTL值由源端设备设置。当报文中的TTL降为0时，报文会被丢弃。同时，丢弃报文的设备会根据报文头中的源IP地址向源端发送ICMP错误消息。

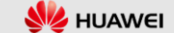


目的端的网络层在接收并处理报文以后，需要决定下一步对报文该做如何处理。IP报文头中的协议字段标识了将会继续处理报文的协议。与以太帧头中的Type字段类似，协议字段也是一个十六进制数。该字段可以标识网络层协议，如ICMP（Internet Control Message Protocol，因特网控制报文协议），也可以标识上层协议，如TCP（Transmission Control Protocol，传输控制协议，对应值0x06）、UDP（User Datagram Protocol，用户数据报协议，对应值0x11）。



总结

- 子网掩码的作用是什么？
- IP报文头部中TTL字段的作用是什么？
- 网关的作用是什么？



1. 32位的IP子网掩码用于区分IP地址中的网络号和主机号。网络号表示网络或子网，主机号表示网络或子网中的主机。
2. 如果网络中存在环路，则IP报文可能会在网络中循环而无法到达目的端。TTL字段限定了IP报文的生存时间，保证无法到达目的端的报文最终被丢弃。
3. 网关是指接收并处理本地网段主机发送的报文并转发到目的网段的设备。


谢谢
www.huawei.com


ICMP协议
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

Internet控制报文协议ICMP（Internet Control Message Protocol）是网络层的一个重要协议。ICMP协议用来在网络设备间传递各种差错和控制信息，它对于收集各种网络信息、诊断和排除各种网络故障具有至关重要的作用。使用基于ICMP的应用时，需要对ICMP的工作原理非常熟悉。

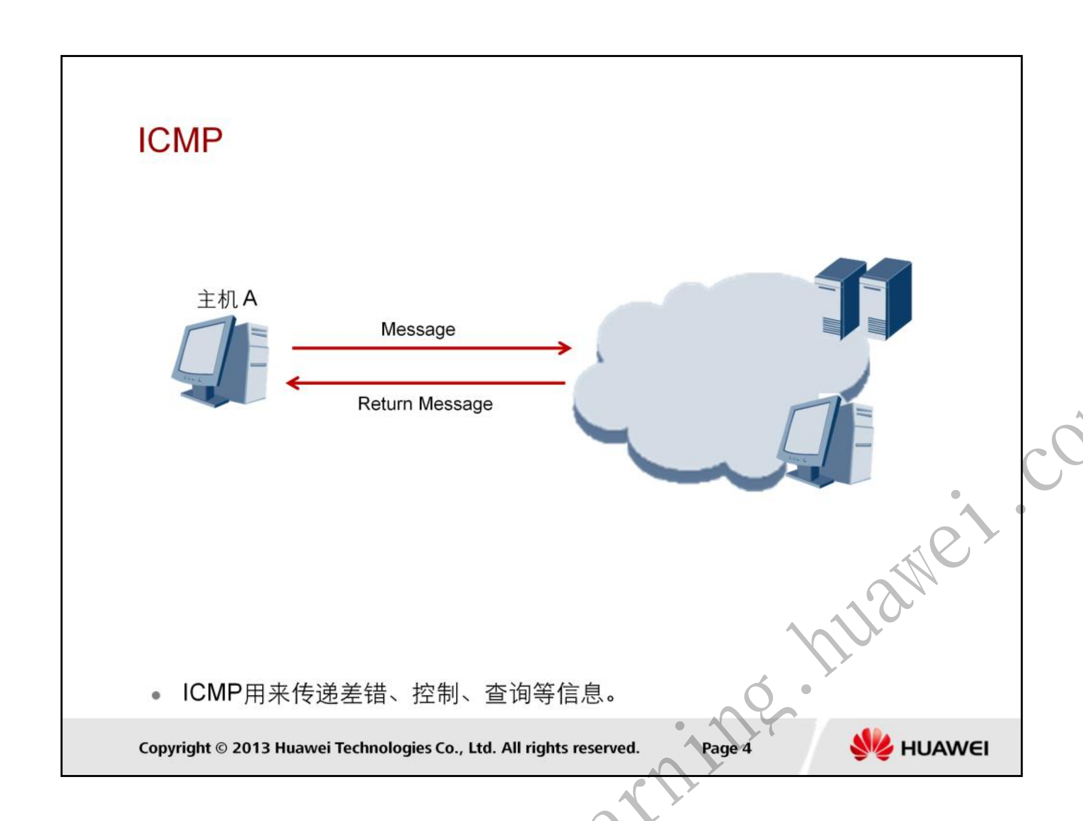




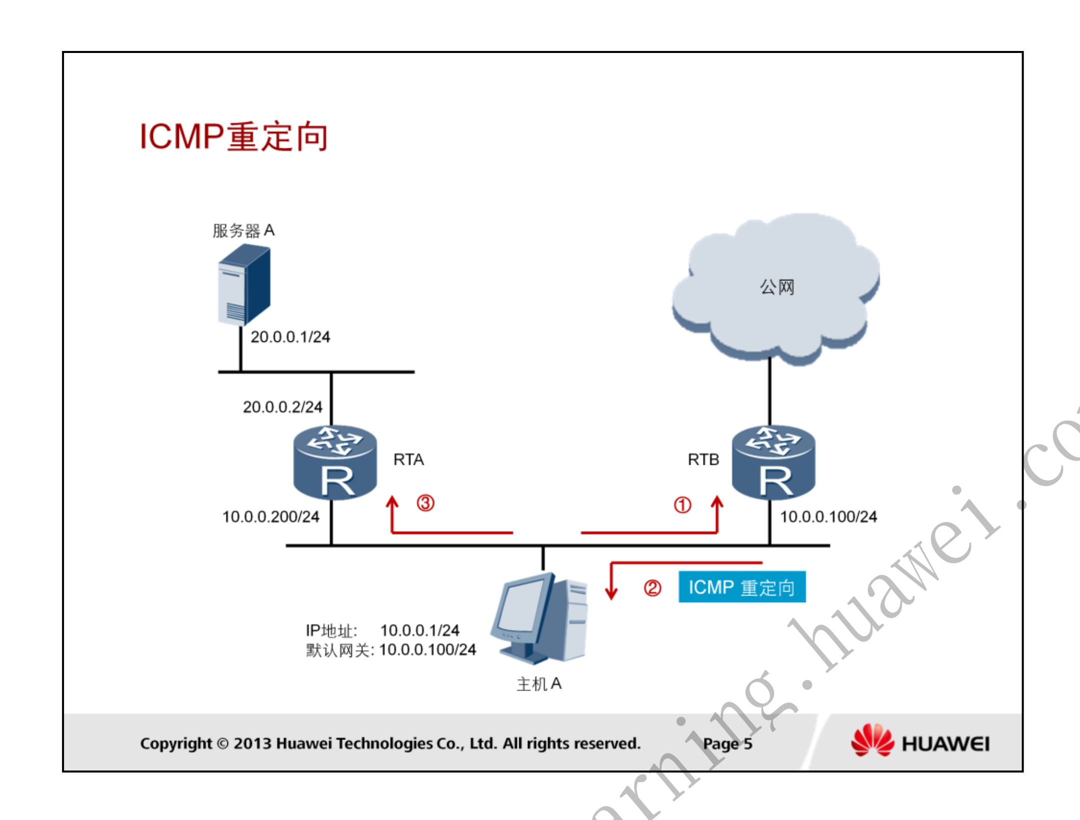
学习目标

学完本课程后，您应该能：

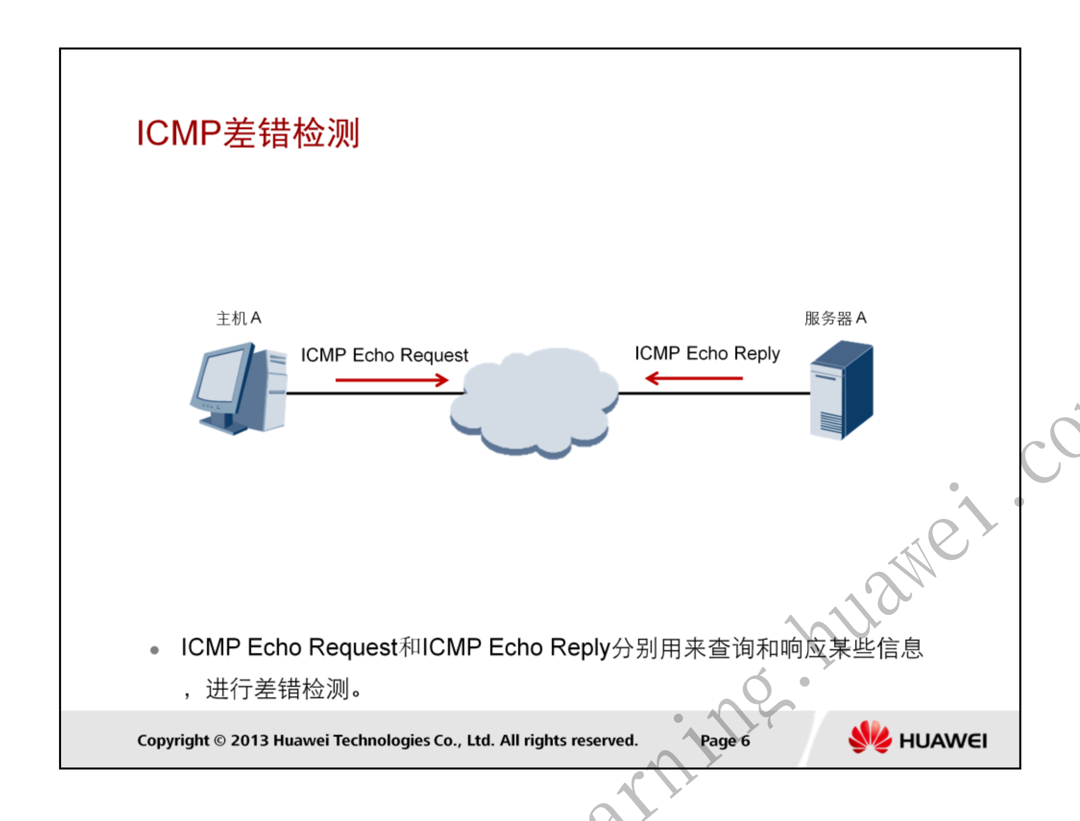
- 描述ICMP的应用场景
- 理解常见的ICMP报文类型
- 掌握Ping和Tracert的应用



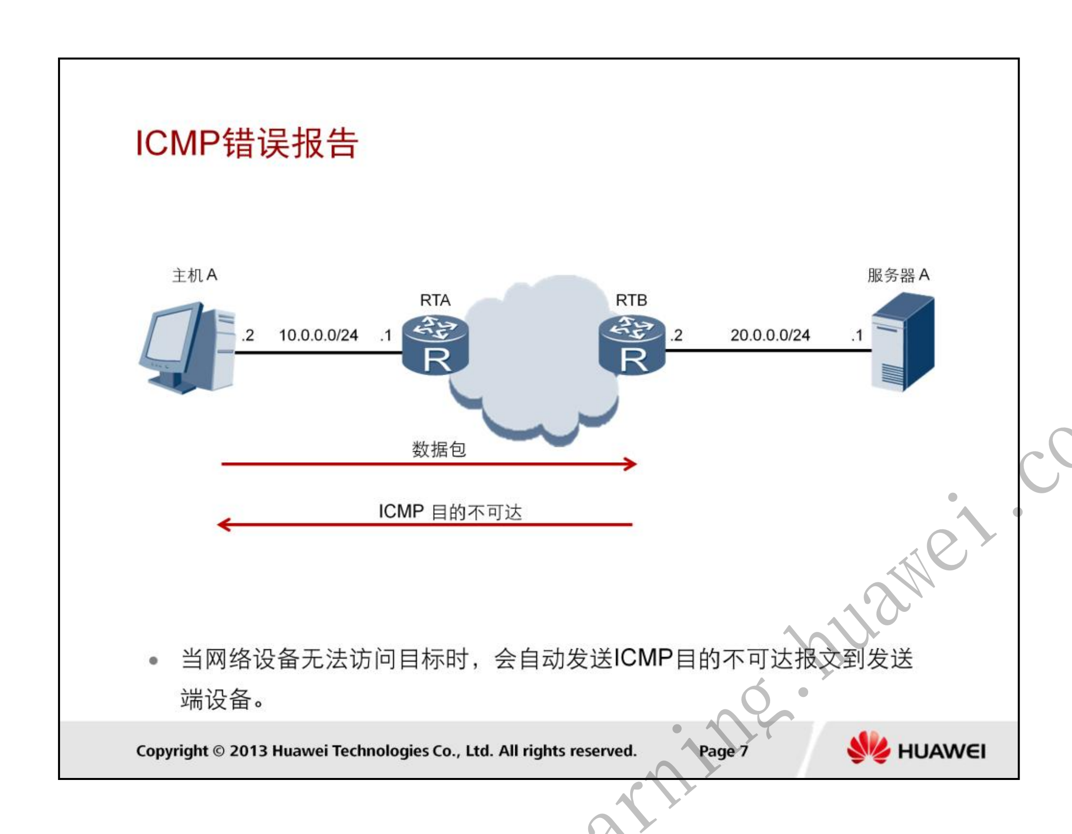
ICMP是TCP/IP协议簇的核心协议之一，它用于在IP网络设备之间发送控制报文，传递差错、控制、查询等信息。



ICMP Redirect重定向消息用于支持路由功能。如图所示，主机A希望发送报文到服务器A，于是根据配置的默认网关地址向网关RTB发送报文。网关RTB收到报文后，检查报文信息，发现报文应该转发到与源主机在同一网段的另一个网关设备RTA，因为此转发路径是更优的路径。所以RTB会向主机发送一个Redirect消息，通知主机直接向另一个网关RTA发送该报文。主机收到Redirect消息后，向RTA发送报文，RTA会将报文转发给服务器A。



ICMP Echo消息常用于诊断源和目的之间的网络连通性，还可以提供其他信息，如报文往返时间等。



ICMP定义了各种错误消息，用于诊断网络连接性问题；根据这些错误消息，源设备可以判断出数据传输失败的原因。比如，如果网络中发生了环路，导致报文在网络中循环，最终TTL超时，这种情况下网络设备会发送TTL超时消息给发送端设备。又比如如果目的不可达，则中间的网络设备会发送目的不可达消息给发送端设备。目的不可达的情况有多种，如果是网络设备无法找到目的网络，则发送目的网络不可达消息；如果网络设备无法找到目的网络中的目的主机，则发送目的主机不可达消息。

ICMP数据包格式



- Type表示ICMP消息类型，Code表示同一消息类型中的不同信息。

 Page 8 

ICMP消息封装在IP报文中。ICMP消息的格式取决于Type和Code字段，其中Type字段为消息类型，Code字段包含该消息类型的具体参数。后面的校验和字段用于检查消息是否完整。消息中包含32比特的可变参数，这个字段一般不使用，通常设置为0。在ICMP Redirect消息中，这个字段用来指定网关IP地址，主机根据这个地址将报文重定向到指定网关。在Echo请求消息中，这个字段包含标识符和序号，源端根据这两个参数将收到的回复消息与本端发送的Echo请求消息进行关联。尤其是当源端向目的端发送了多个Echo请求消息时，需要根据标识符和序号将Echo请求和回复进行一一对应。

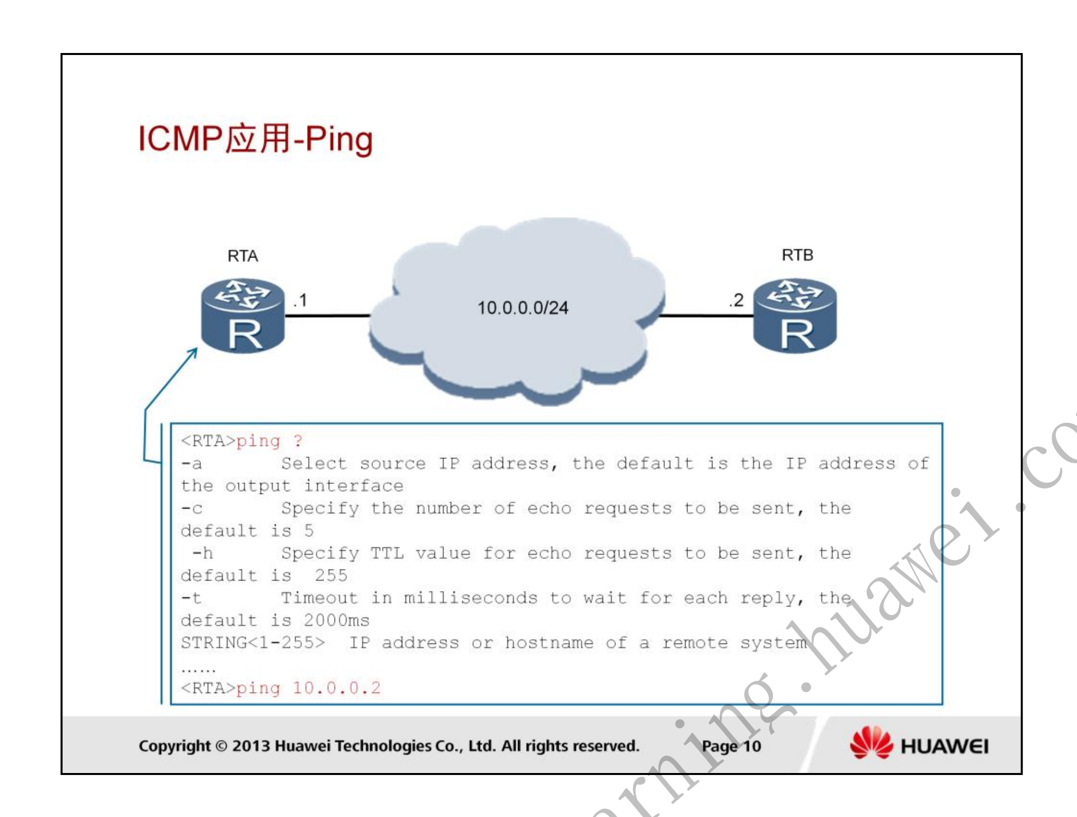
ICMP消息类型和编码类型

类型	编码	描述
0	0	Echo Reply
3	0	网络不可达
3	1	主机不可达
3	2	协议不可达
3	3	端口不可达
5	0	重定向
8	0	Echo Request



ICMP定义了多种消息类型，用于不同的场景。有些消息不需要Code字段来描述具体类型参数，仅用Type字段表示消息类型。比如，ICMP Echo回复消息的Type字段设置为0。

有些ICMP消息使用Type字段定义消息大类，用Code字段表示消息的具体类型。比如，类型为3的消息表示目的不可达，不同的Code值表示不可达的原因，包括目的网络不可达(Code=0)、目的主机不可达(Code=1)、协议不可达（Code=2）、目的TCP/UDP端口不可达(Code=3) 等。



ICMP的一个典型应用是Ping。Ping是检测网络连通性的常用工具，同时也能够收集其他相关信息。用户可以在Ping命令中指定不同参数，如ICMP报文长度、发送的ICMP报文个数、等待回复响应的超时时间等，设备根据配置的参数来构造并发送ICMP报文，进行Ping测试。

Ping常用的配置参数说明如下：

1. -a *source-ip-address*指定发送ICMP ECHO-REQUEST报文的源IP地址。如果不指定源IP地址，将采用出接口的IP地址作为ICMP ECHO-REQUEST报文发送的源地址。

2. -c *count*指定发送ICMP ECHO-REQUEST报文次数。缺省情况下发送5个ICMP ECHO-REQUEST报文。

3. -h *ttl-value*指定TTL的值。缺省值是255。

4. -t *timeout*指定发送完ICMP ECHO-REQUEST后，等待ICMP ECHO-REPLY的超时时间。

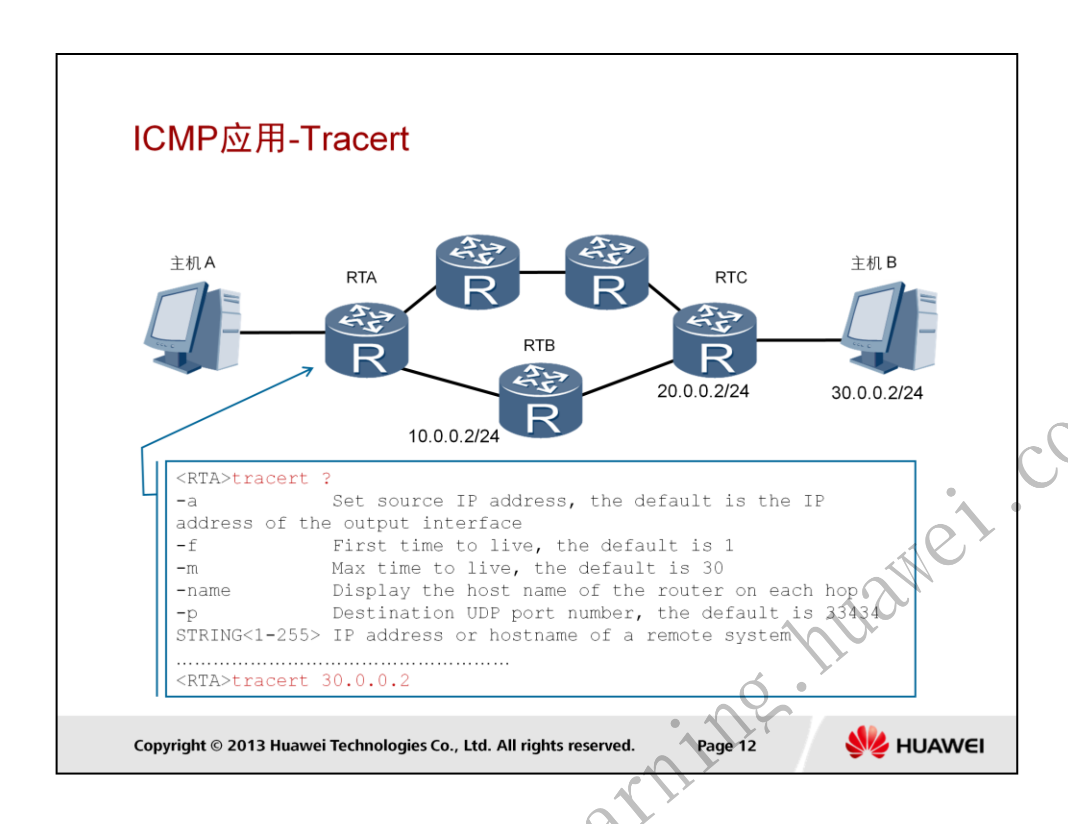
ICMP应用-Ping

```
[RTA]ping 10.0.0.2
  PING 10.0.0.2 : 56  data bytes, press CTRL_C to break
    Reply from 10.0.0.2 : bytes=56 Sequence=1 ttl=255 time=340 ms
    Reply from 10.0.0.2 : bytes=56 Sequence=2 ttl=255 time=10 ms
    Reply from 10.0.0.2 : bytes=56 Sequence=3 ttl=255 time=30 ms
    Reply from 10.0.0.2 : bytes=56 Sequence=4 ttl=255 time=30 ms
    Reply from 10.0.0.2 : bytes=56 Sequence=5 ttl=255 time=30 ms

  --- 10.0.0.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 10/88/340 ms
```

  HUAWEI

Ping命令的输出信息中包括目的地址、ICMP报文长度、序号、TTL值、以及往返时间。序号是包含在Echo回复消息（Type=0）中的可变参数字段，TTL和往返时间包含在消息的IP头中。



ICMP的另一个典型应用是Tracert。Tracert基于报文头中的TTL值来逐跳跟踪报文的转发路径。为了跟踪到达某特定目的地址的路径，源端首先将报文的TTL值设置为1。该报文到达第一个节点后，TTL超时，于是该节点向源端发送TTL超时消息，消息中携带时间戳。然后源端将报文的TTL值设置为2，报文到达第二个节点后超时，该节点同样返回TTL超时消息，以此类推，直到报文到达目的地。这样，源端根据返回的报文中的信息可以跟踪到报文经过的每一个节点，并根据时间戳信息计算往返时间。Tracert是检测网络丢包及时延的有效手段，同时可以帮助管理员发现网络中的路由环路。

Tracert常用的配置参数说明如下：

1. -a *source-ip-address*指定tracert报文的源地址。
2. -f *first-ttl*指定初始TTL。缺省值是1。
3. -m *max-ttl*指定最大TTL。缺省值是30。
4. -name使能显示每一跳的主机名。
5. -p *port*指定目的主机的UDP端口号。

ICMP应用-Tracert

主机A

RTA

RTB

RTC

主机B

10.0.0.2/24

20.0.0.2/24

30.0.0.2/24

```
<RTA>tracert 30.0.0.2
Tracert to 30.0.0.2(30.0.0.2), max hops:30, packet length:40,
press CTRL_C to break
 1 10.0.0.2 130 ms  50 ms  40 ms
 2 20.0.0.2 80 ms  60 ms  80 ms
 3 30.0.0.2 80 ms  60 ms  70 ms
```

- Tracert显示数据包在网络传输过程中所经过的每一跳。

 Page 13 HUAWEI

源端（RTA）向目的端（主机B）发送一个UDP报文，TTL值为1，目的UDP端口号是大于30000的一个数，因为在大多数情况下，大于30000的UDP端口号是任何一个应用程序都不可能使用的端口号。

第一跳（RTB）收到源端发出的UDP报文后，判断出报文的目的IP地址不是本机IP地址，将TTL值减1后，判断出TTL值等于0，则丢弃报文并向源端发送一个ICMP超时（Time Exceeded）报文（该报文中含有第一跳的IP地址10.0.0.2），这样源端就得到了RTB的地址。

源端收到RTB的ICMP超时报文后，再次向目的端发送一个UDP报文，TTL值为2。

第二跳（RTC）收到源端发出的UDP报文后，回应一个ICMP超时报文，这样源端就得到了RTC的地址（20.0.0.2）。

以上过程不断进行，直到目的端收到源端发送的UDP报文后，判断出目的IP地址是本机IP地址，则处理此报文。根据报文中的目的UDP端口号寻找占用此端口号的上层协议，因目的端没有应用程序使用该UDP端口号，则向源端返回一个ICMP端口不可达（Destination Unreachable）报文。

源端收到ICMP端口不可达报文后，判断出UDP报文已经到达目的端，则停止Tracert程序，从而得到数据报文从源端到目的端所经历的路径（10.0.0.2；20.0.0.2；30.0.0.2）。



总结

- Ping使用到的是哪两类ICMP消息?
- 当网络设备收到TTL值为0的IP报文时，会如何操作?

 Page 14 

1. Ping利用ICMP Echo请求消息（Type值为8）来发起检测目的可达性。目的端收到ICMP Echo请求消息后，根据IP报文头中的源地址向源端发送ICMP Echo回复消息（Type值为0）。
2. 如果IP数据报在到达目的之前TTL值已经降为0，则收到IP数据包的网络设备会丢弃该数据包，并向源端发送ICMP消息通知源端TTL超时。


谢谢
www.huawei.com


ARP协议
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

网络设备有数据要发送给另一台网络设备时，必须要知道对方的网络层地址（即IP地址）。IP地址由网络层来提供，但是仅有IP地址是不够的，IP数据报文必须封装成帧才能通过数据链路进行发送。数据帧必须要包含目的MAC地址，因此发送端还必须获取到目的MAC地址。通过目的IP地址而获取目的MAC地址的过程是由ARP（Address Resolution Protocol）协议来实现的。

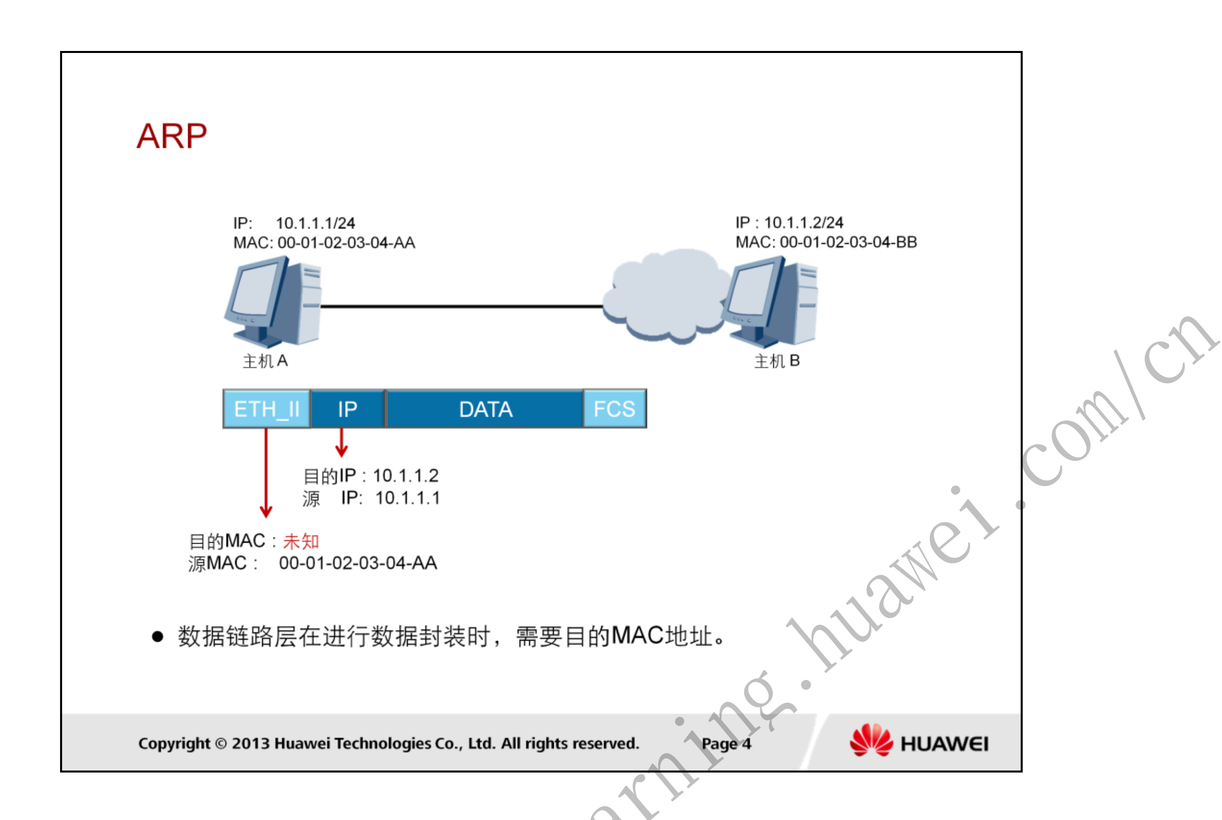




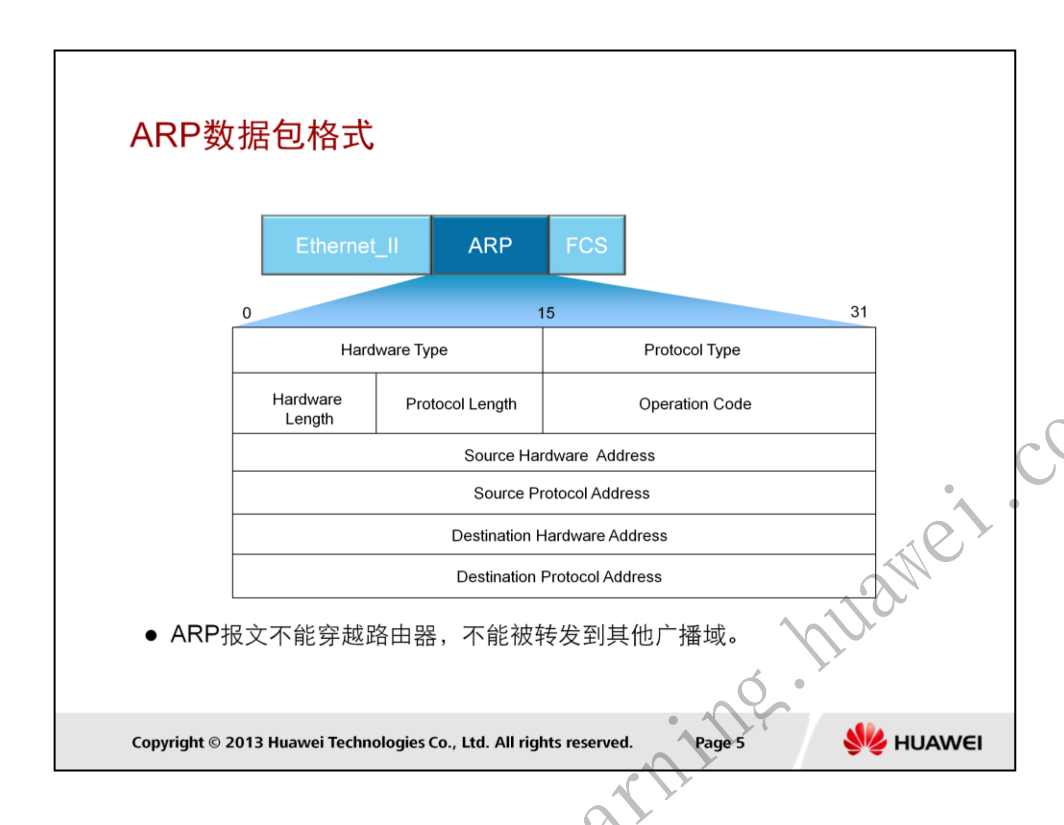
学习目标

学完本课程后，您应该能：

- 掌握ARP的工作原理
- 理解ARP缓存表的作用

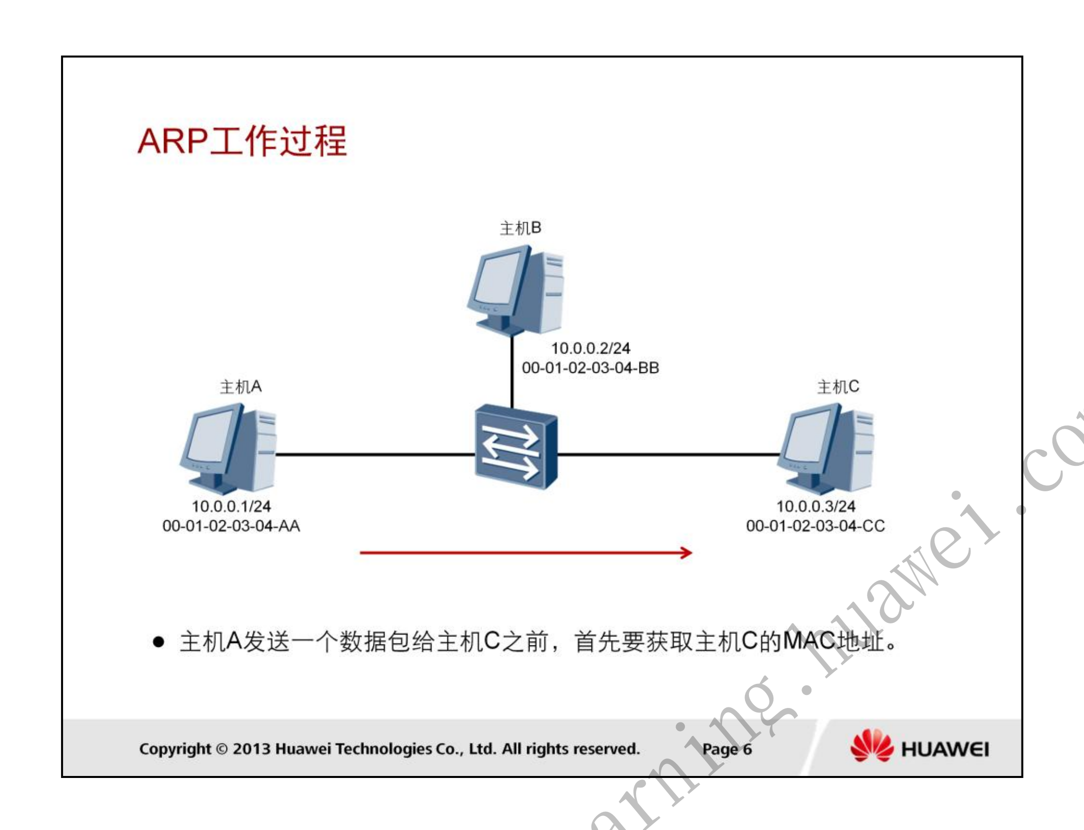


一台网络设备要发送数据给另一太网络设备时，必须要知道对方的IP地址。但是，仅有IP地址是不够的，因为IP数据报文必须封装成帧才能通过数据链路进行发送，而数据帧必须要包含目的MAC地址，因此发送端还必须获取到目的MAC地址。每一个网络设备在数据封装前都需要获取下一跳的MAC地址。IP地址由网络层来提供，MAC地址通过ARP协议来获取。ARP协议是TCP/IP协议簇中的重要组成部分，ARP能够通过目的IP地址发现目标设备的MAC地址，从而实现数据链路层的可达性。

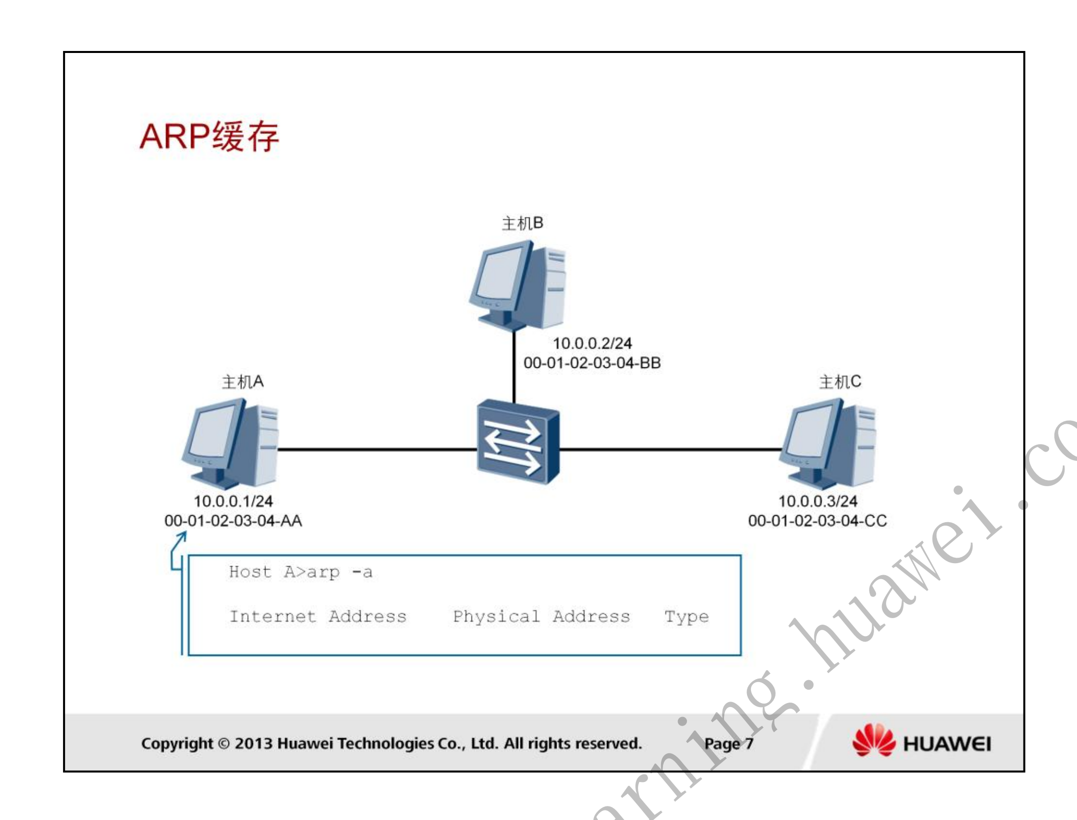


网络设备通过ARP报文来发现目的MAC地址。ARP报文中包含以下字段：

1. Hardware Type表示硬件地址类型，一般为以太网；
2. Protocol Type表示三层协议地址类型，一般为IP；
3. Hardware Length和Protocol Length为MAC地址和IP地址的长度，单位是字节；
4. Operation Code指定了ARP报文的类型，包括ARP request和ARP reply；
5. Source Hardware Address指的是发送ARP报文的设备MAC地址；
6. Source Protocol Address指的是发送ARP报文的设备IP地址；
7. Destination Hardware Address指的是接收者MAC地址，在ARP request报文中，该字段值为0；
8. Destination Protocol Address指的是指接收者的IP地址。

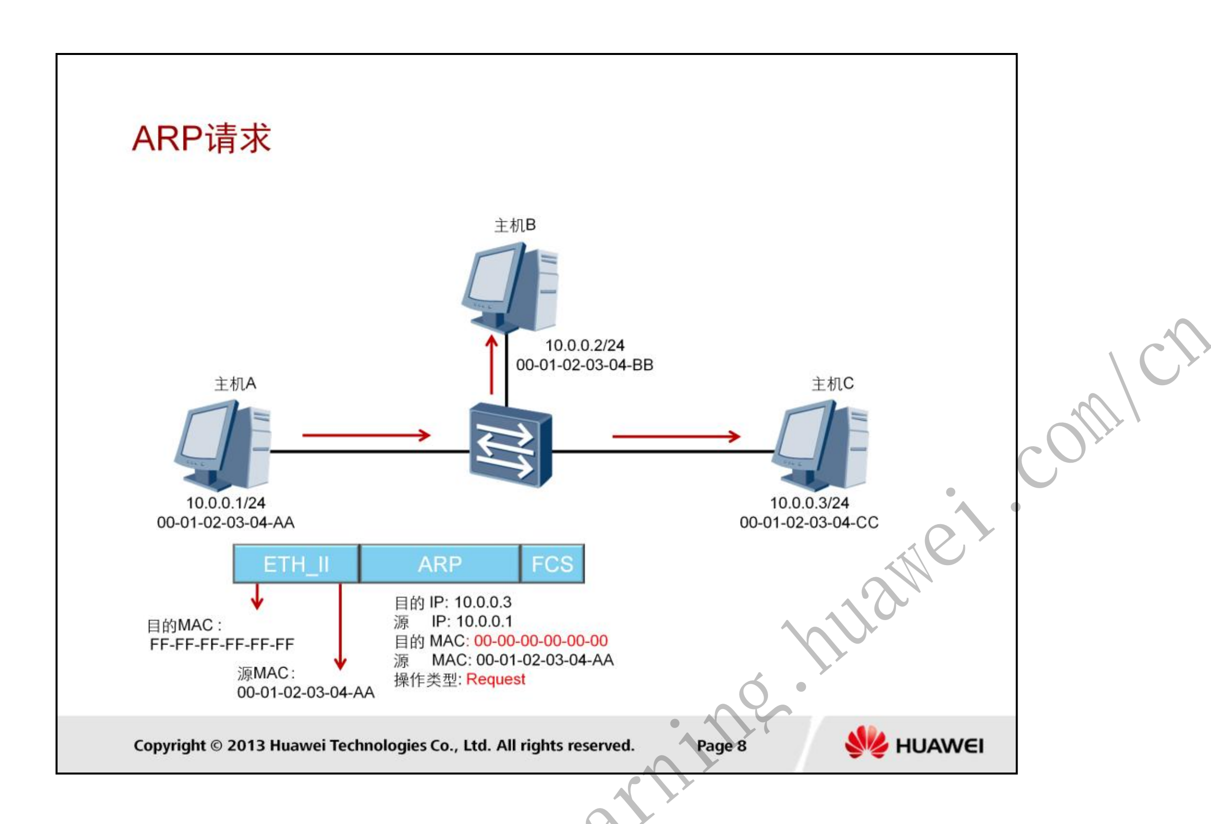


通过ARP协议，网络设备可以建立目标IP地址和MAC地址之间的映射。网络设备通过网络层获取到目的IP地址之后，还要判断目的MAC地址是否已知。

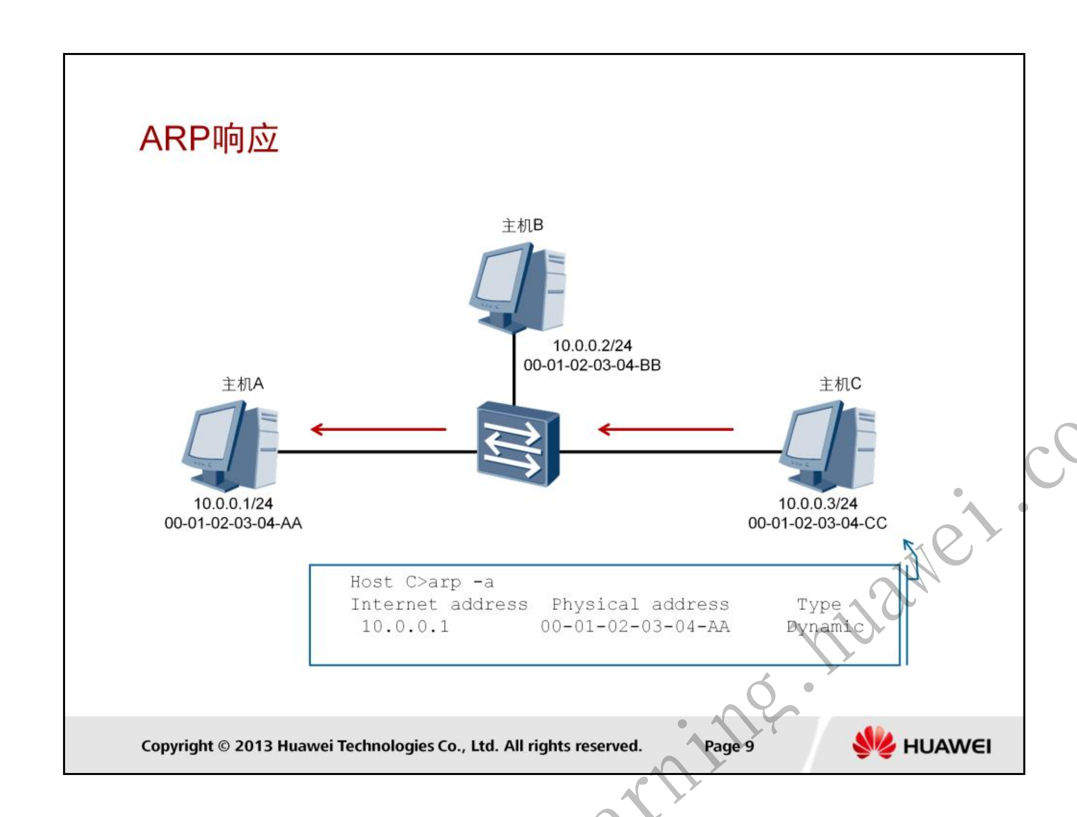


网络设备一般都有一个ARP缓存（ARP Cache），ARP缓存用来存放IP地址和MAC地址的关联信息。在发送数据前，设备会先查找ARP缓存表。如果缓存表中存在对方设备的MAC地址，则直接采用该MAC地址来封装帧，然后将帧发送出去。如果缓存表中不存在相应信息，则通过发送ARP Request报文来获得它。学习到的IP地址和MAC地址的映射关系会被放入ARP缓存表中存放一段时间。在有效期内，设备可以直接从这个表中查找目的MAC地址来进行数据封装，而无需进行ARP查询。过了这段有效期，ARP表项会被自动删除。

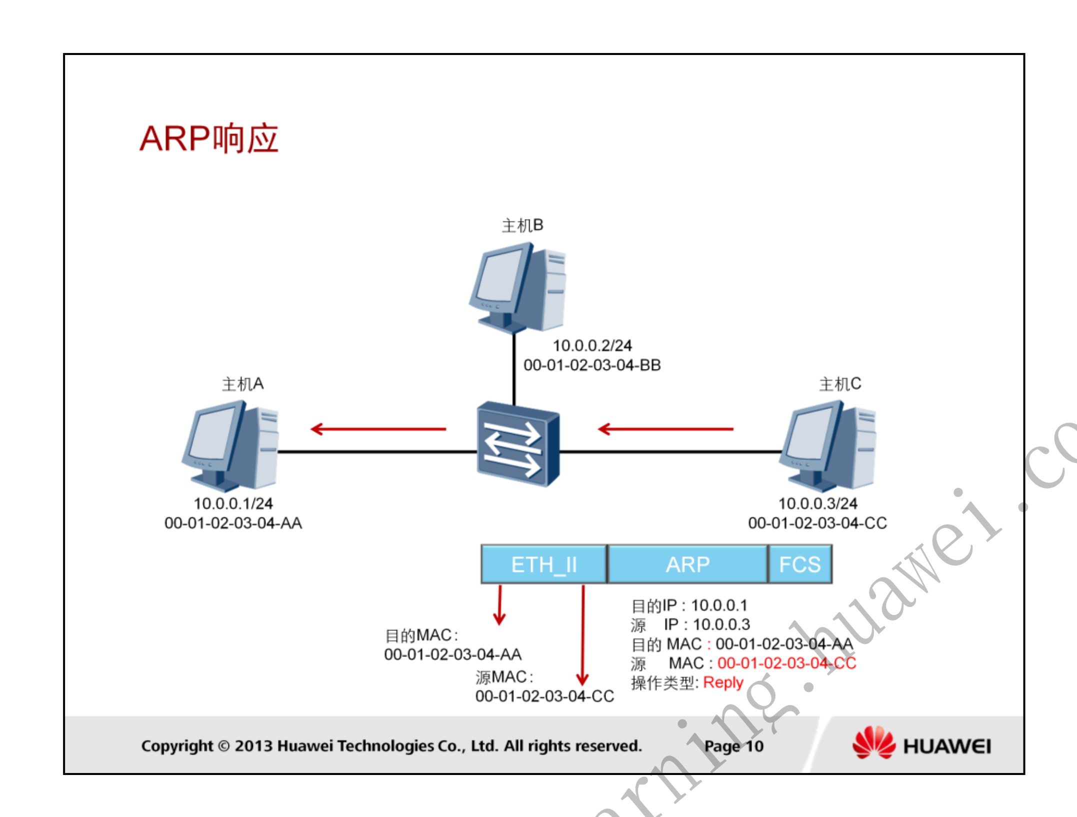
如果目标设备位于其他网络，则源设备会在ARP缓存表中查找网关的MAC地址，然后将数据发送给网关，网关再把数据转发给目的设备。



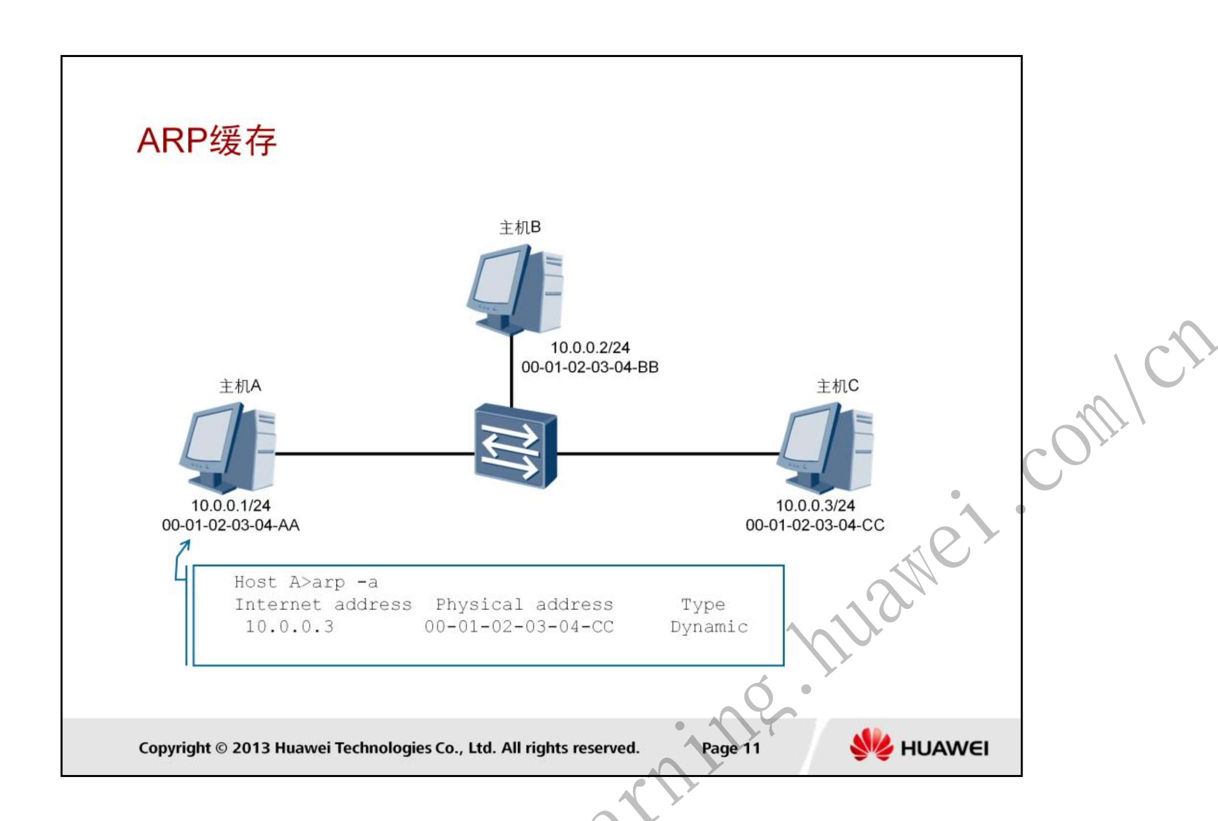
本例中，主机A的ARP缓存表中不存在主机C的MAC地址，所以主机A会发送ARP Request来获取目的MAC地址。ARP Request报文封装在以太帧里。帧头中的源MAC地址为发送端主机A的MAC地址。此时，由于主机A不知道主机C的MAC地址，所以目的MAC地址为广播地址FF-FF-FF-FF-FF-FF。ARP Request报文中包含源IP地址、目的IP地址、源MAC地址、目的MAC地址，其中目的MAC地址的值为0。ARP Request报文会在整个网络上传播，该网络中所有主机包括网关都会接收到此ARP Request报文。网关将会阻止该报文发送到其他网络上。



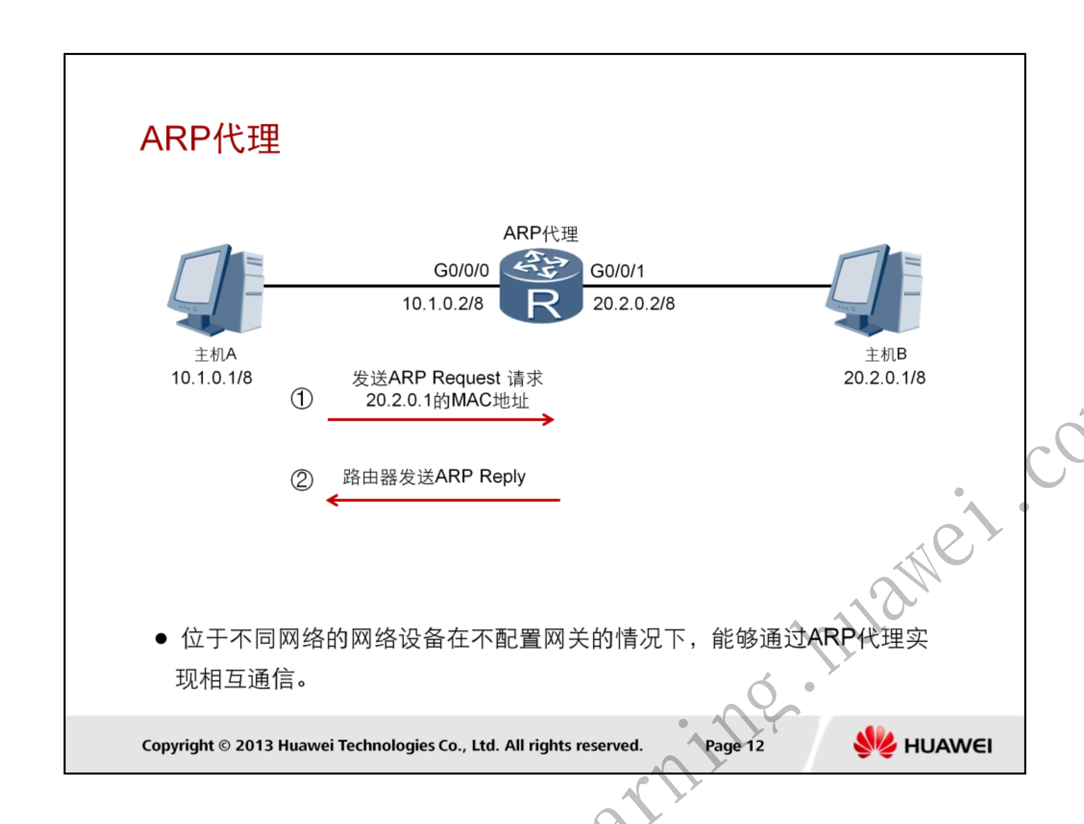
所有的主机接收到该ARP Request报文后，会检查它的目的协议地址字段与自身的IP地址是否匹配。如果不匹配，则该主机将不会响应该ARP Request报文。如果匹配，则该主机会将ARP报文中的源MAC地址和源IP地址信息记录到自己的ARP缓存表中，然后通过ARP Reply报文进行响应。



主机C会向主机A回应ARP Reply报文。ARP Reply报文中的源协议地址是主机C自己的IP地址，目标协议地址是主机A的IP地址，目的MAC地址是主机A的MAC地址，源MAC地址是自己的MAC地址，同时Operation Code被设置为Reply。ARP Reply报文通过单播传送。

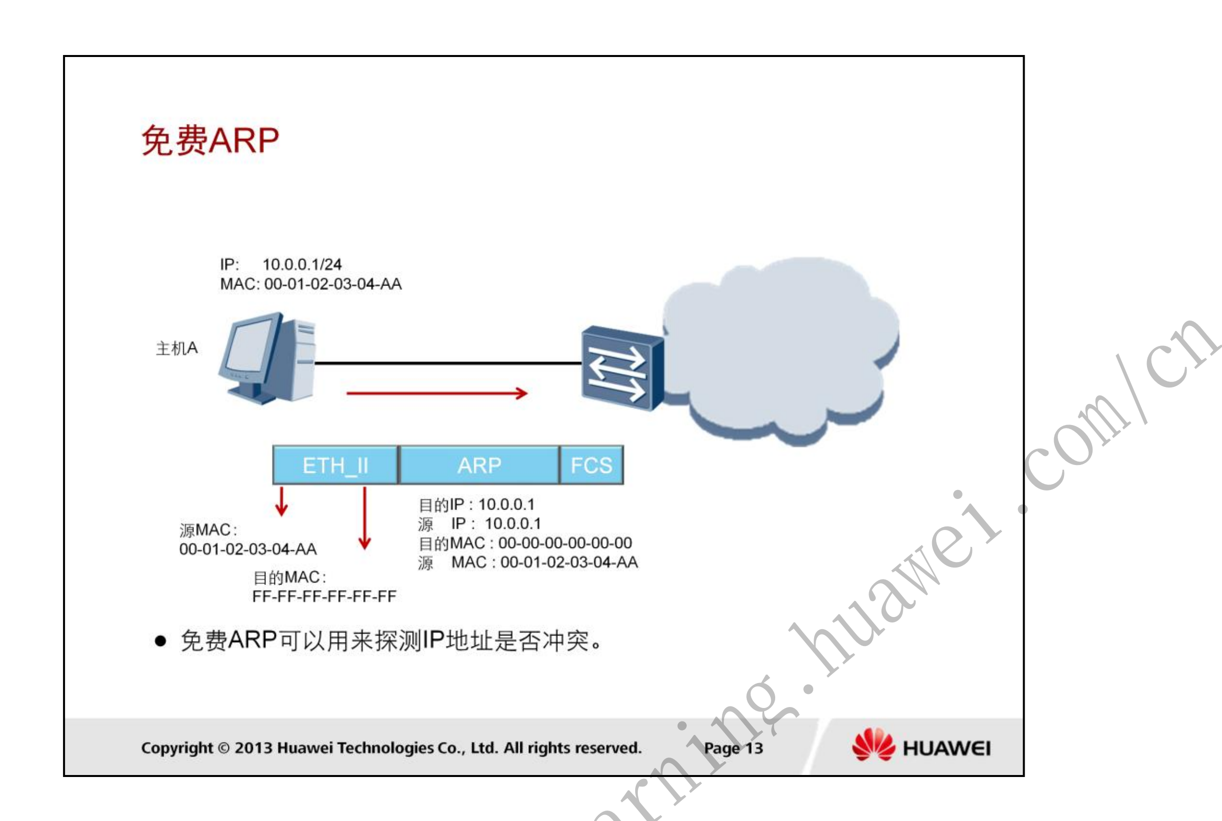


主机A收到ARP Reply以后，会检查ARP报文中目的MAC地址是否与自己的MAC匹配。如果匹配，ARP报文中的源MAC地址和源IP地址会被记录到主机A的ARP缓存表中。ARP表项的老化超时时间缺省为1200秒。



在上述例子的组网中，主机A需要与主机B通信时，目的IP地址与本机的IP地址位于不同网络，但是由于主机A未配置网关，所以它将会以广播形式发送ARP Request报文，请求主机B的MAC地址。但是，广播报文无法被路由器转发，所以主机B无法收到主机A的ARP请求报文，当然也就无法应答。

在路由器上启用代理ARP功能，就可以解决这个问题。启用代理ARP后，路由器收到这样的请求，会查找路由表，如果存在主机B的路由表项，路由器将会使用自己的G0/0/0接口的MAC地址来回应该ARP Request。主机A收到ARP Reply后，将以路由器的G0/0/0接口MAC地址作为目的MAC地址进行数据转发。



主机被分配了IP地址或者IP地址发生变更后，必须立刻检测其所分配的IP地址在网络上是否是唯一的，以避免地址冲突。主机通过发送ARP request报文来进行地址冲突检测。

主机A将ARP Request广播报文中的目的IP地址字段设置为自己的IP地址，该网络中所有主机包括网关都会接收到此报文。当目的IP地址已经被某一个主机或网关使用时，该主机或网关就会回应ARP reply报文。通过这种方式，主机A就能探测到IP地址冲突了。



总结

- 网络设备在什么情况下会发送ARP request?
- 网络设备什么时候会产生免费ARP?



1. 源设备在发送数据给目的设备前，会首先查看自身的ARP缓存，查找ARP缓存中是否存在目的设备的IP地址和MAC地址的映射。如果存在则直接使用，如果不存在则会发送ARP request。
2. 当网络上的一个设备被分配了IP地址或者IP地址发生变更后，可以通过免费ARP来检查IP地址是否冲突。

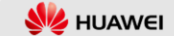

谢谢
www.huawei.com


传输层协议
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

传输层定义了主机应用程序之间端到端的连通性。传输层中最为常见的两个协议分别是传输控制协议TCP（Transmission Control Protocol ）和 用户数据包协议UDP（User Datagram Protocol）。

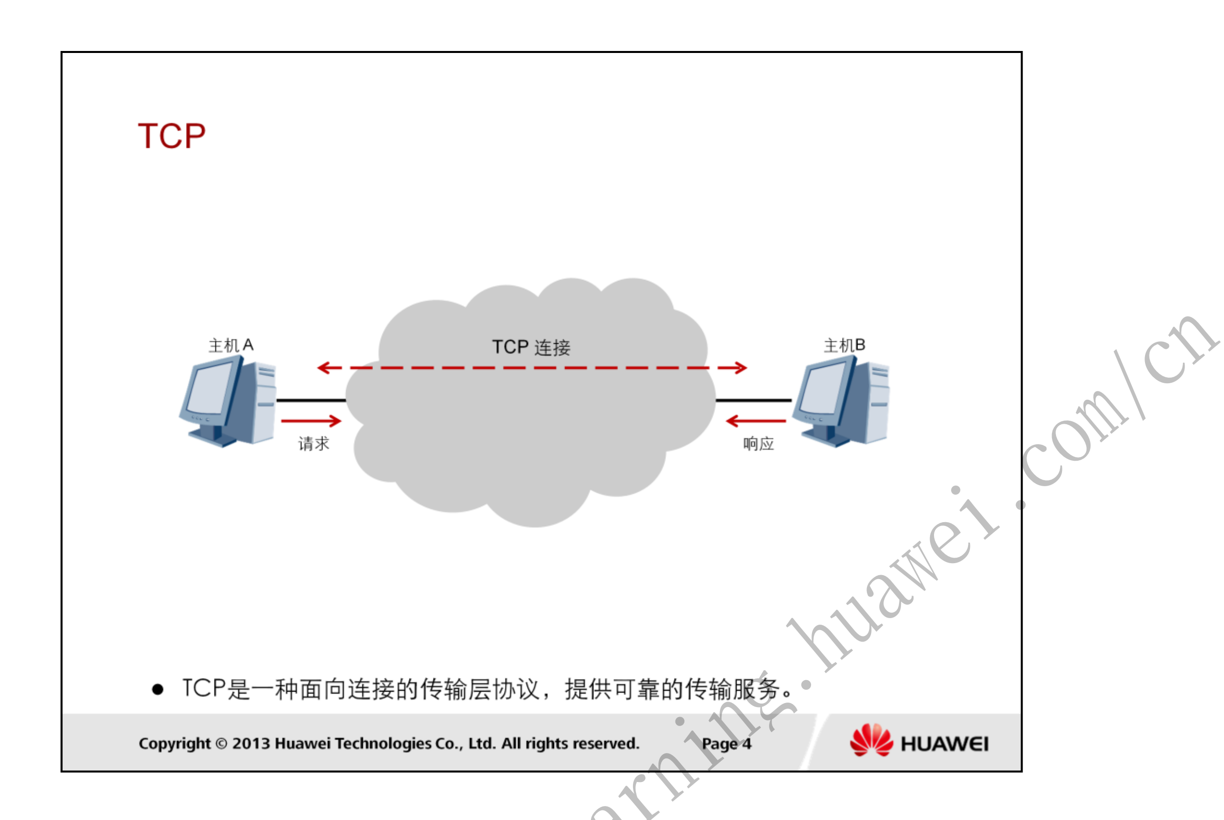




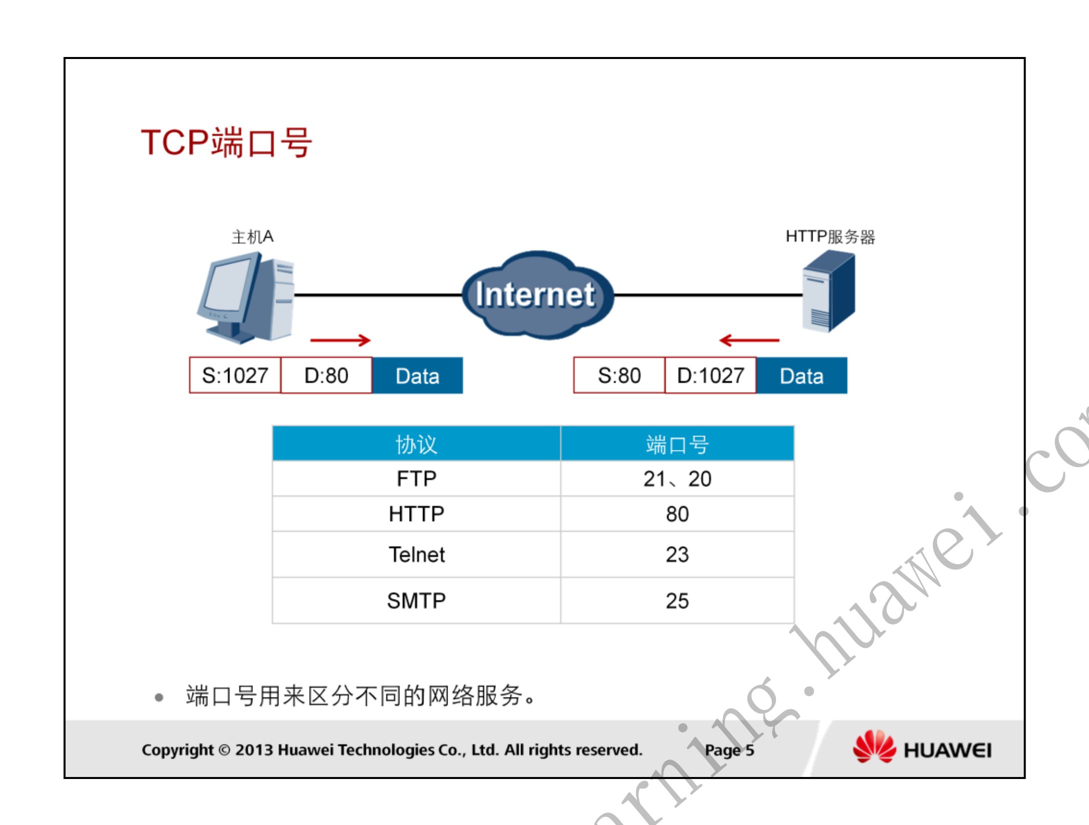
学习目标

学完本课程后，您应该能：

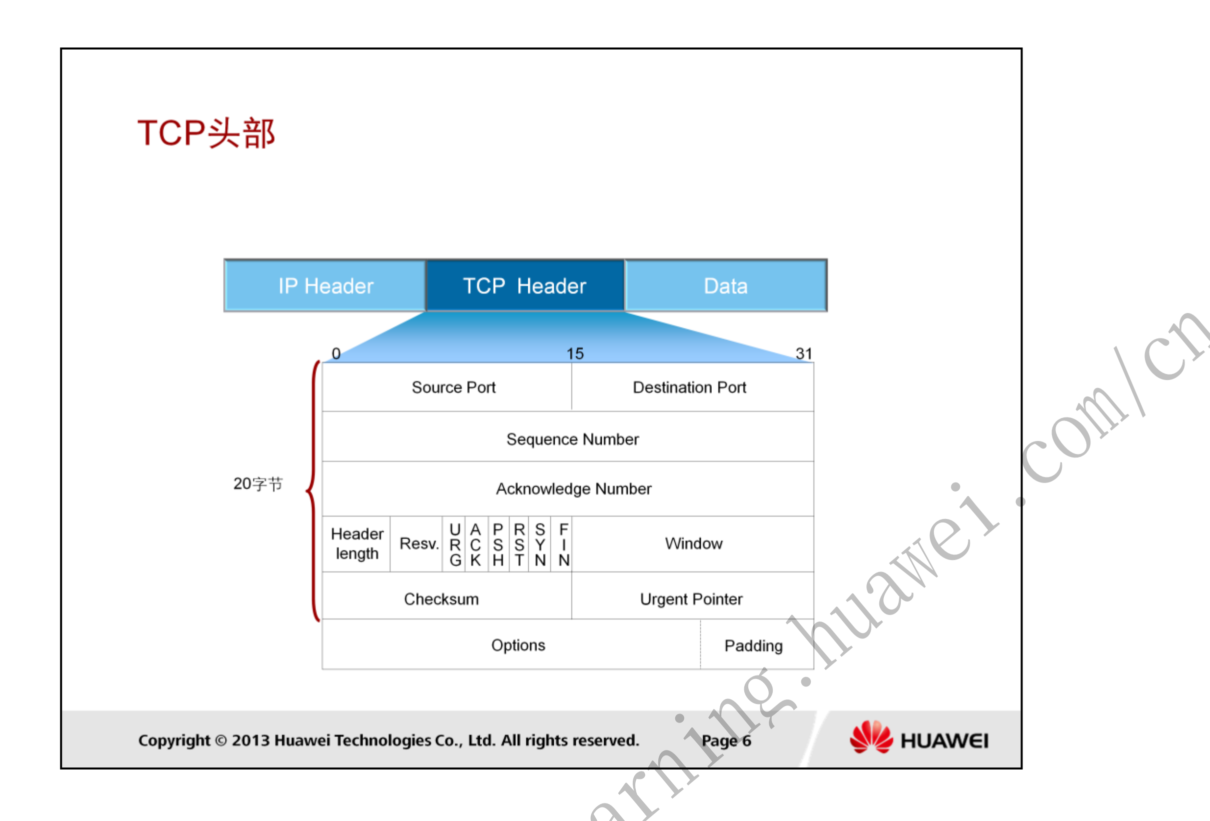
- 掌握TCP和UDP的工作原理
- 描述TCP和UDP的报文格式
- 了解常见服务的应用端口号



TCP位于TCP/IP模型的传输层，它是一种面向连接的端到端协议。TCP作为传输控制协议，可以为主机提供可靠的数据传输。TCP需要依赖网络协议为主机提供可用的传输路径。在本例中，两台主机在通信之前，需要TCP在它们之间建立可靠的传输通道。



TCP允许一个主机同时运行多个应用进程。每台主机可以拥有多个应用端口，每对端口号、源和目标IP地址的组合唯一地标识了一个会话。端口分为知名端口和动态端口。有些网络服务会使用固定的端口，这类端口称为知名端口，端口号范围为0-1023。如FTP、HTTP、Telnet、SNMP服务均使用知名端口。动态端口号范围从1024到65535，这些端口号一般不固定分配给某个服务，也就是说许多服务都可以使用这些端口。只要运行的程序向系统提出访问网络的申请，那么系统就可以从这些端口号中分配一个供该程序使用。

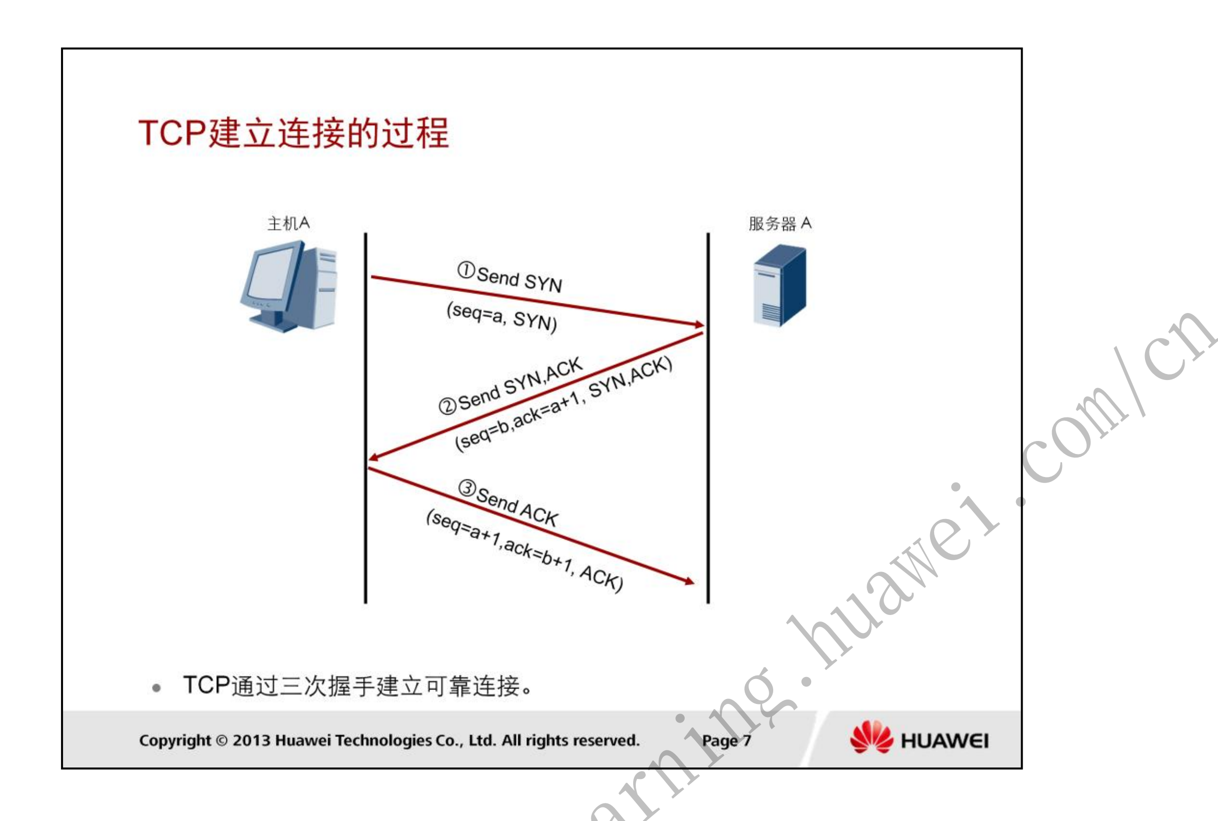


TCP通常使用IP作为网络层协议,这时TCP数据段被封装在IP数据包内。

TCP数据段由TCP Header（头部）和TCP Data（数据）组成。TCP最多可以有60个字节的头部，如果没有Options字段，正常的长度是20字节。

TCP Header是由如上图标识一些字段组成，这里列出几个常用字段。

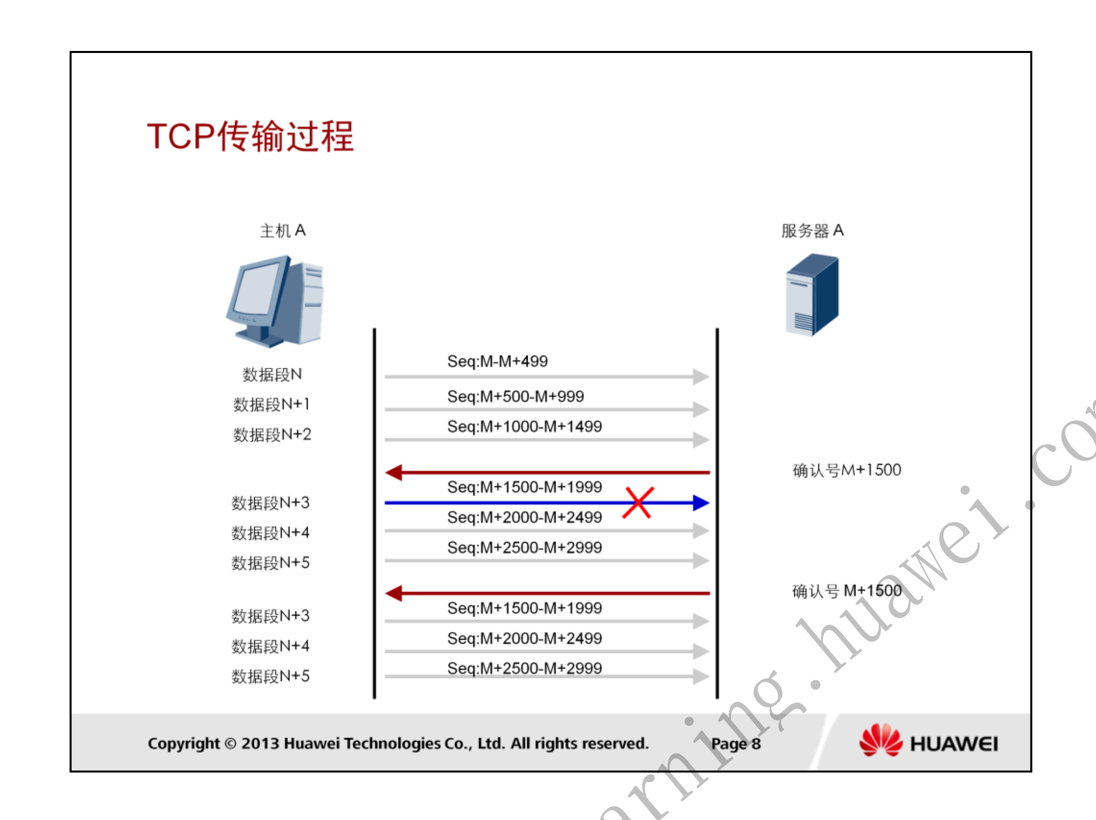
1. 16位源端口号：源主机的应用程序使用的端口号。
2. 16位目的端口号：目的主机的应用程序使用的端口号。每个TCP头部都包含源和目的端的端口号，这两个值加上IP头部中的源IP地址和目的IP地址可以唯一确定一个TCP连接。
3. 32位序列号：用于标识从发送端发出的不同的TCP数据段的序号。数据段在网络中传输时，它们的顺序可能会发生变化；接收端依据此序列号，便可按照正确的顺序重组数据。
4. 32位确认序列号：用于标识接收端确认收到的数据段。确认序列号为成功收到的数据序列号加1。
5. 4位头部长度：表示头部占32bit字的数目，它能表达的TCP头部最大长度为60字节。
6. 16位窗口大小：表示接收端期望通过单次确认而收到的数据的大小。由于该字段为16位，所以窗口大小的最大值为65535字节，该机制通常用来进行流量控制。
7. 16位校验和：校验整个TCP报文段，包括TCP头部和TCP数据。该值由发送端计算和记录并由接收端进行验证。



TCP是一种可靠的，面向连接的全双工传输层协议。

TCP连接的建立是一个三次握手的过程。如图所示：

1. 主机A（通常也称为客户端）发送一个标识了SYN的数据段，表示期望与服务器A建立连接，此数据段的序列号（seq）为a。
2. 服务器A回复标识了SYN+ACK的数据段，此数据段的序列号（seq）为b，确认序列号为主机A的序列号加1（a+1），以此作为对主机A的SYN报文的确认。
3. 主机A发送一个标识了ACK的数据段，此数据段的序列号（seq）为a+1，确认序列号为服务器A的序列号加1（b+1），以此作为对服务器A的SYN报文段的确认。

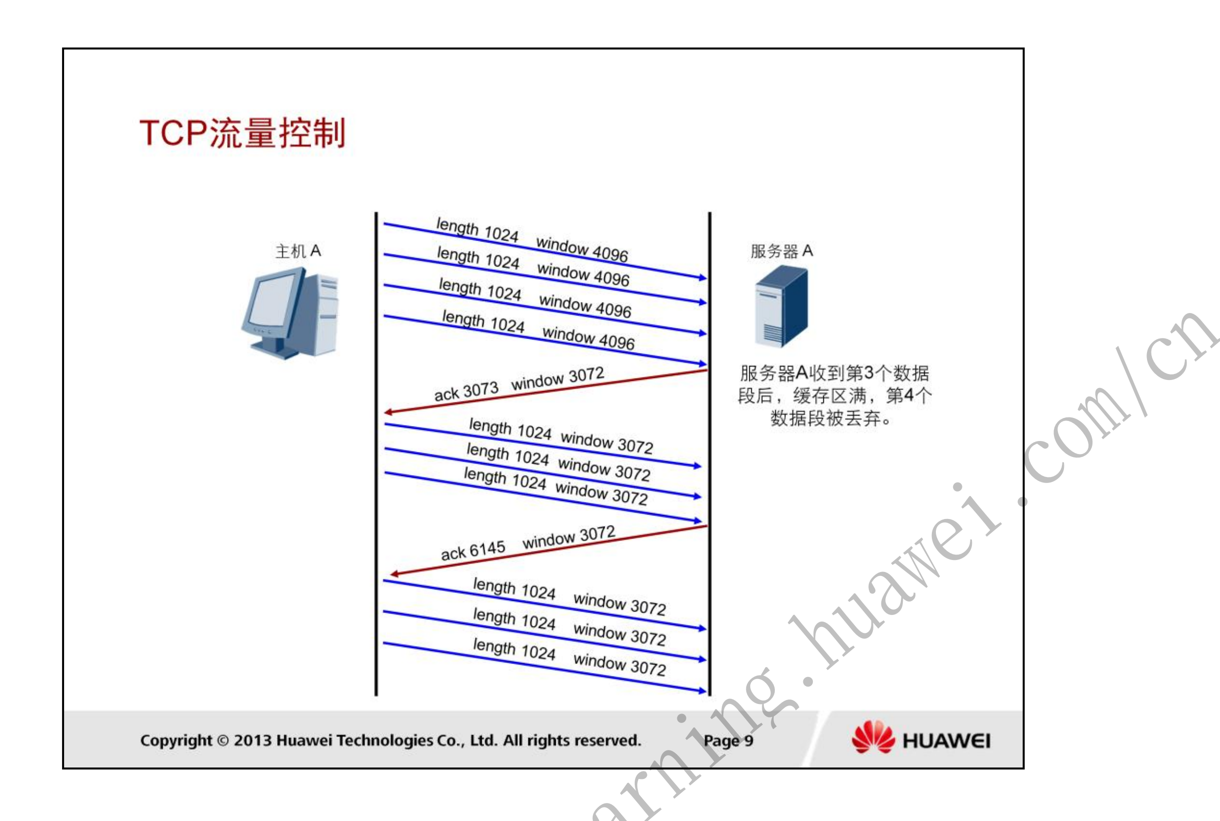


TCP的可靠传输还体现在TCP使用了确认技术来确保目的设备收到了从源设备发来的数据，并且是准确无误的。

确认技术的工作原理如下：

目的设备接收到源设备发送的数据段时，会向源端发送确认报文，源设备收到确认报文后，继续发送数据段，如此重复。

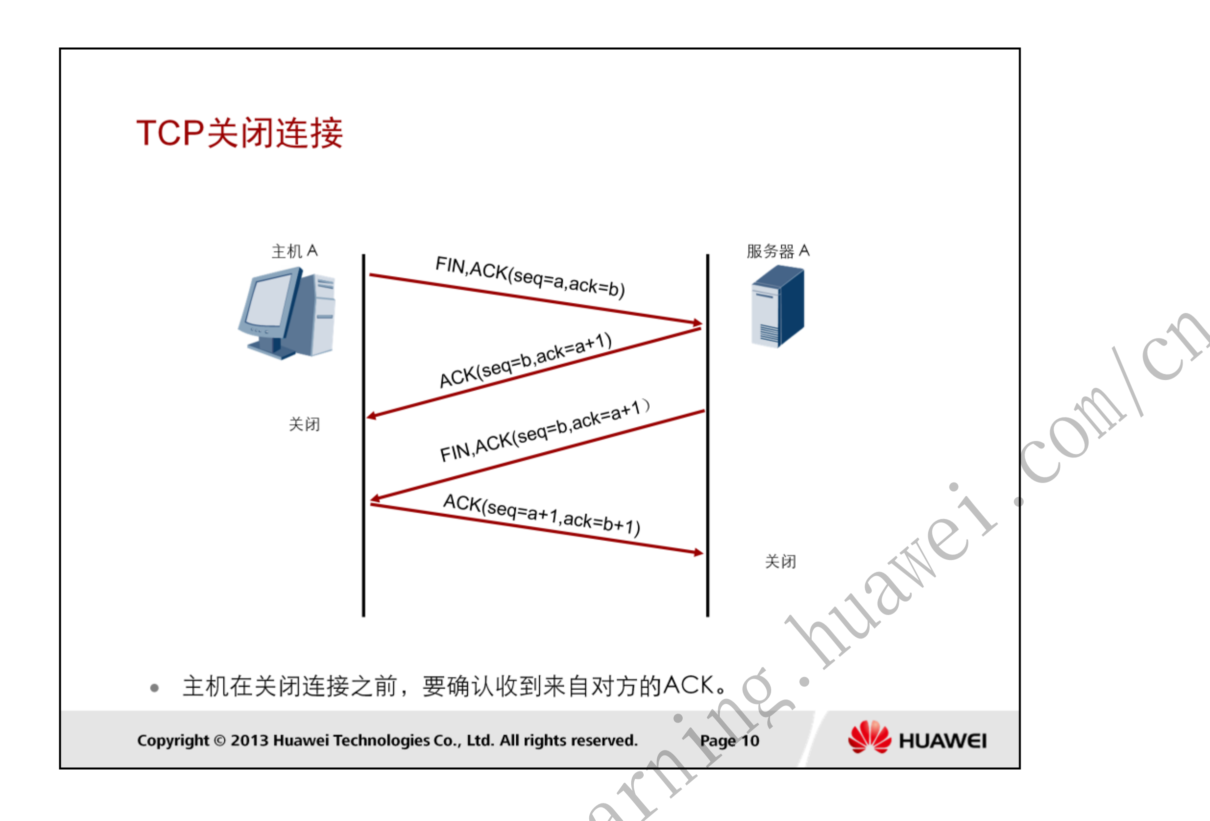
如图所示，主机A向服务器A发送TCP数据段，为描述方便假定每个数据段的长度都是500个字节。当服务器A成功收到序列号是M+1499的字节以及之前的所有字节时，会以序列号M+1499+1=M+1500进行确认。另外，由于数据段N+3传输失败，所以服务器A未能收到序列号为M+1500的字节，因此服务器A还会再次以序列号M+1500进行确认。



TCP滑动窗口技术通过动态改变窗口大小来实现对端到端设备之间的数据传输进行流量控制。

如图所示，主机A和服务器A之间通过滑动窗口来实现流量控制。为方便理解，此例中只考虑主机A发送数据给服务器A时，服务器A通过滑动窗口进行的流量控制。

主机A向服务器发送4个长度为1024字节的数据段，其中主机的窗口大小为4096个字节。服务器A收到第3个数据段后，缓存区满，第4个数据段被丢弃。服务器以ACK3073响应，窗口大小调整为3072，表明服务器的缓冲区只能处理3072个字节的数据段。于是主机A改变其发送速率，发送窗口大小为3072的数据段。



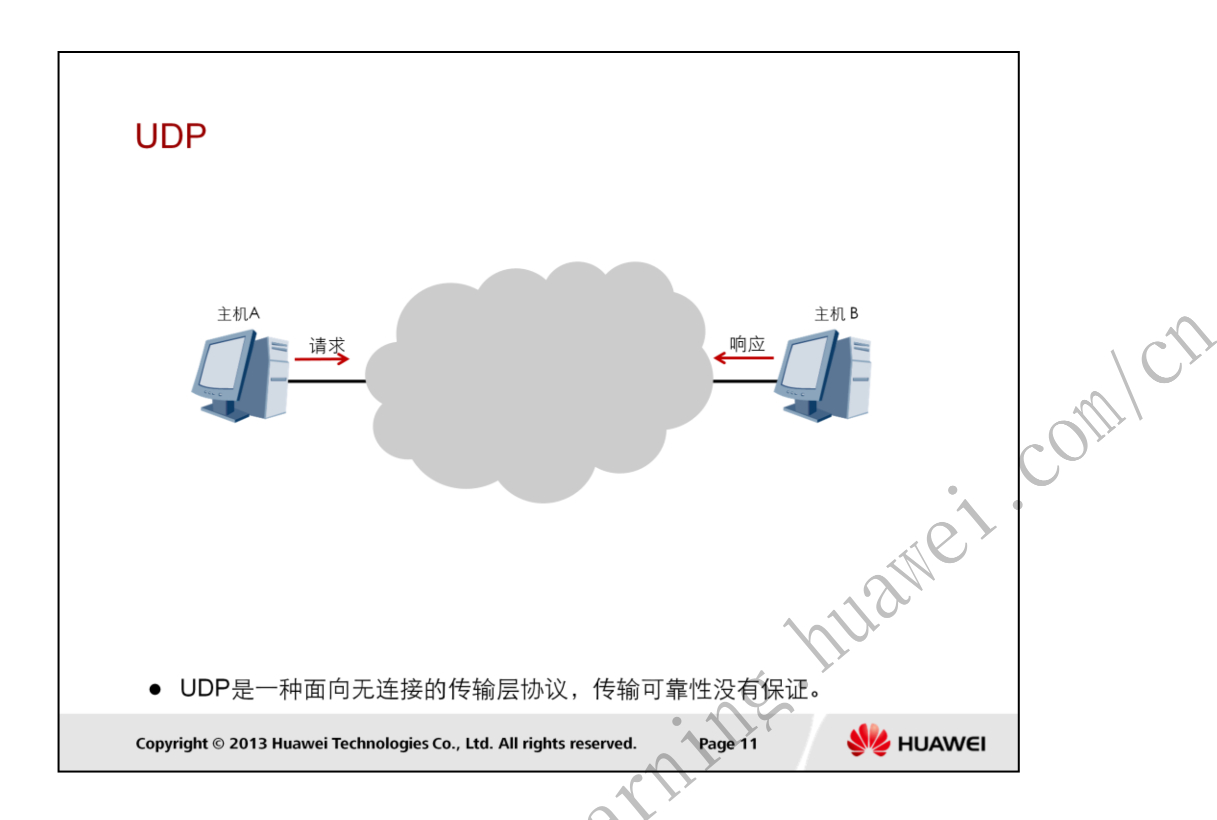
TCP支持全双工模式传输数据，这意味着同一时刻两个方向都可以进行数据的传输。在传输数据之前，TCP通过三次握手建立的实际上是两个方向的连接，因此在传输完毕后，两个方向的连接必须都关闭。

TCP连接的建立是一个三次握手的过程，而TCP连接的终止则要经过四次握手。

如图所示：

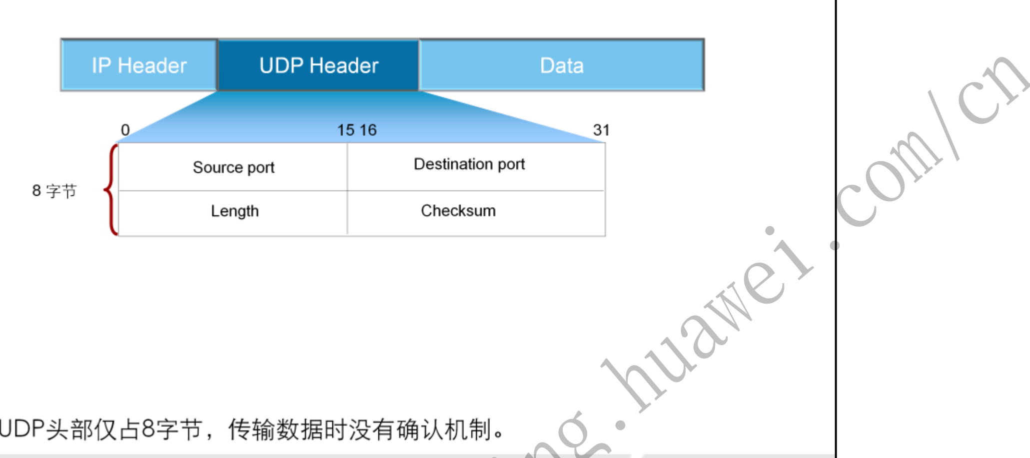
1. 主机A想终止连接，于是发送一个标识了FIN，ACK的数据段，序列号为a，确认序列号为b。
2. 服务器A回应一个标识了ACK的数据段，序列号为b，确认序号为a+1，作为对主机A的FIN报文的确认。
3. 服务器A想终止连接，于是向主机A发送一个标识了FIN，ACK的数据段，序列号为b，确认序列号为a+1。
4. 主机A回应一个标识了ACK的数据段，序列号为a+1，确认序号为b+1，作为对服务器A的FIN报文的确认。

以上四次交互便完成了两个方向连接的关闭。



当应用程序对传输的可靠性要求不高，但是对传输速度和延迟要求较高时，可以用UDP协议来替代TCP协议在传输层控制数据的转发。UDP将数据从源端发送到目的端时，无需事先建立连接。UDP采用了简单、易操作的机制在应用程序间传输数据，没有使用TCP中的确认技术或滑动窗口机制，因此UDP不能保证数据传输的可靠性，也无法避免接收到重复数据的情况。

UDP头部

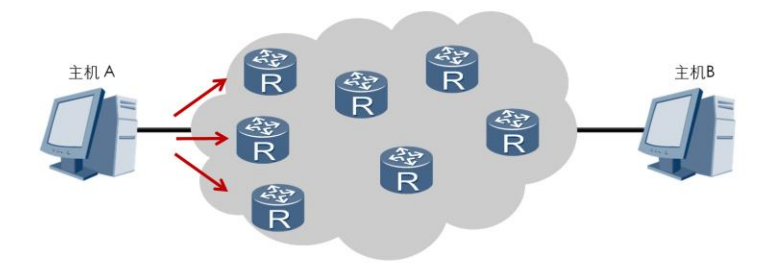


 Page 12 

UDP报文分为UDP报文头和UDP数据区域两部分。报头由源端口、目的端口、报文长度以及校验和组成。UDP适合于实时数据传输，如语音和视频通信。相比于TCP，UDP的传输效率更高、开销更小，但是无法保障数据传输的可靠性。UDP头部的标识如下：

1. 16位源端口号：源主机的应用程序使用的端口号。
2. 16位目的端口号：目的主机的应用程序使用的端口号。
3. 16位UDP长度：是指UDP头部和UDP数据的字节长度。因为UDP头部长度为8字节，所以该字段的最小值为8。
4. 16位UDP校验和：该字段提供了与TCP校验字段同样的功能；该字段是可选的。

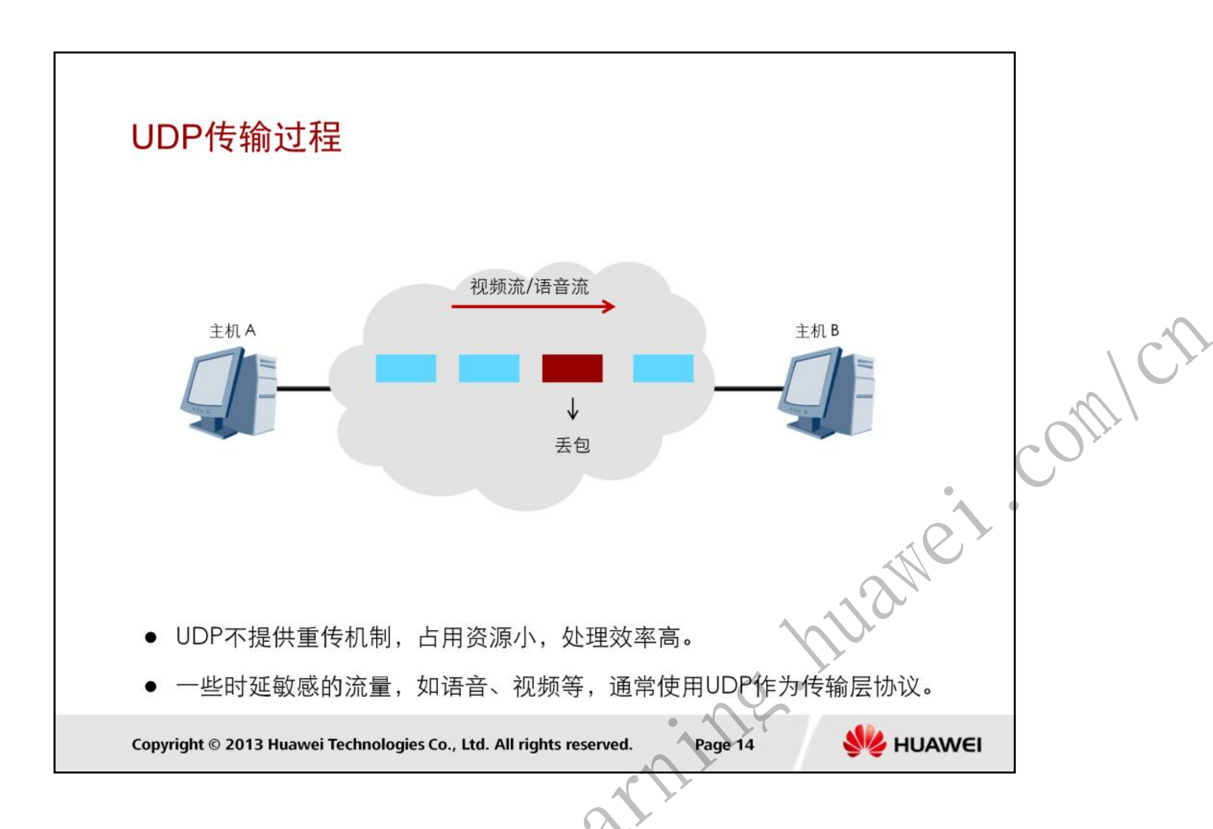
UDP传输过程



- 使用UDP传输数据时，由应用程序根据需要提供报文到达确认、排序、流量控制等功能。



主机A发送数据包时，这些数据包是以有序的方式发送到网络中的，每个数据包独立地在网络中被发送，所以不同的数据包可能会通过不同的网络路径到达主机B。这样的情况下，先发送的数据包不一定先到达主机B。因为UDP数据包没有序号，主机B将无法通过UDP协议将数据包按照原来的顺序重新组合，所以此时需要应用程序提供报文的到达确认、排序和流量控制等功能。通常情况下，UDP采用实时传输机制和时间戳来传输语音和视频数据。



UDP适合传输对延迟敏感的流量，如语音和视频。

在使用TCP协议传输数据时，如果一个数据段丢失或者接收端对某个数据段没有确认，发送端会重新发送该数据段。

TCP重新发送数据会带来传输延迟和重复数据，降低了用户的体验。对于迟延敏感的应用，少量的数据丢失一般可以被忽略，这时使用UDP传输将能够提升用户的体验。



总结

- TCP头部中的确认标识位有什么作用?
- TCP头部中有哪些标识位参与TCP三次握手?



1. TCP报文头中的ACK标志位用于目的端对已收到数据的确认。目的端成功收到序列号为x的字节及之前的所有字节后，会以序列号x+1进行确认。
2. 在TCP的三次握手过程中，要使用SYN和ACK标志位来请求建立连接和确认建立连接。


谢谢
www.huawei.com


数据转发过程
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

TCP/IP协议簇和底层协议配合，保证了数据能够实现端到端的传输。数据传输过程是一个非常复杂的过程，例如数据在转发的过程中会进行一系列的封装和解封装。对于网络工程师来说，只有深入地理解了数据在各种不同设备上的转发过程，才能够对网络进行正确的分析和检测。

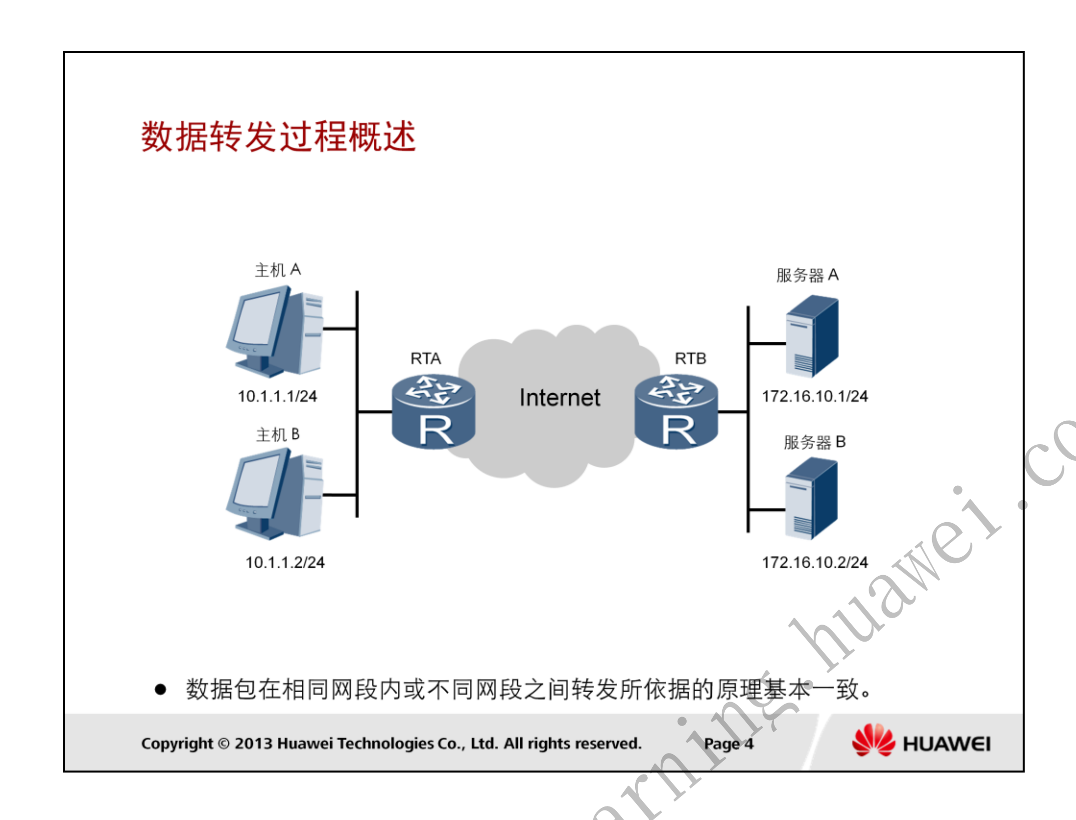




学习目标

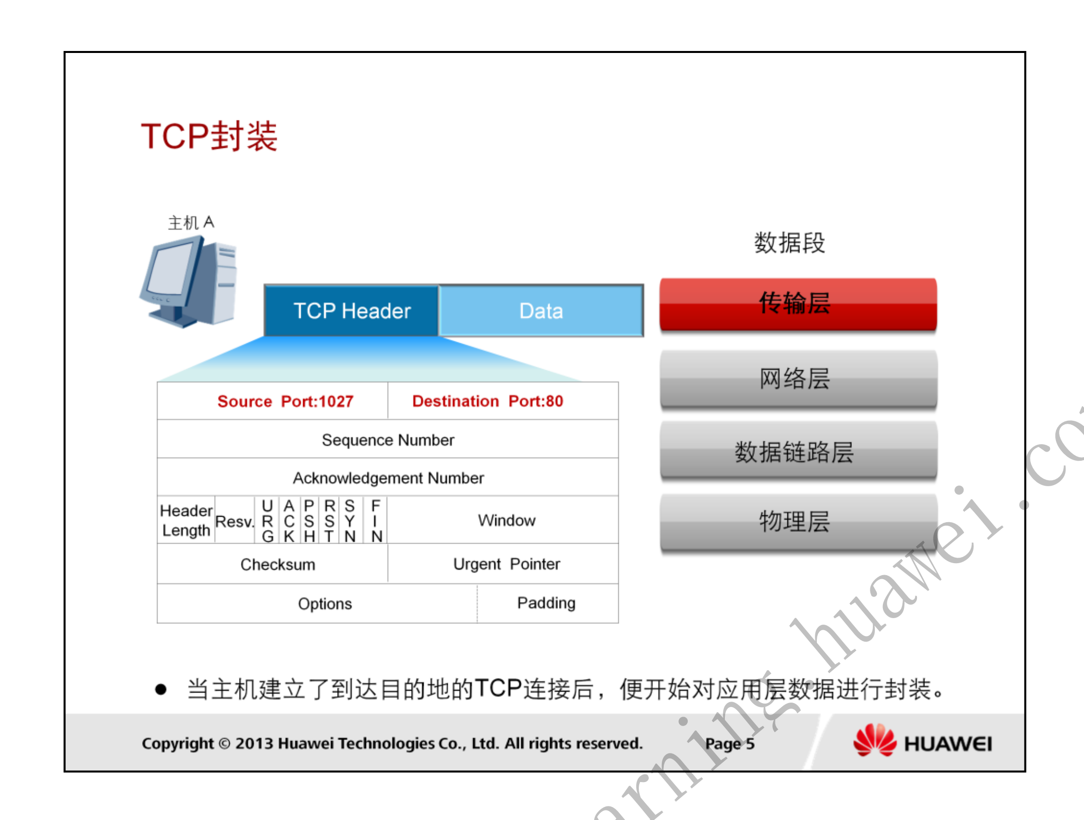
学完本课程后，您应该能：

- 掌握数据封装和解封装的过程
- 处理数据转发过程中的基本故障

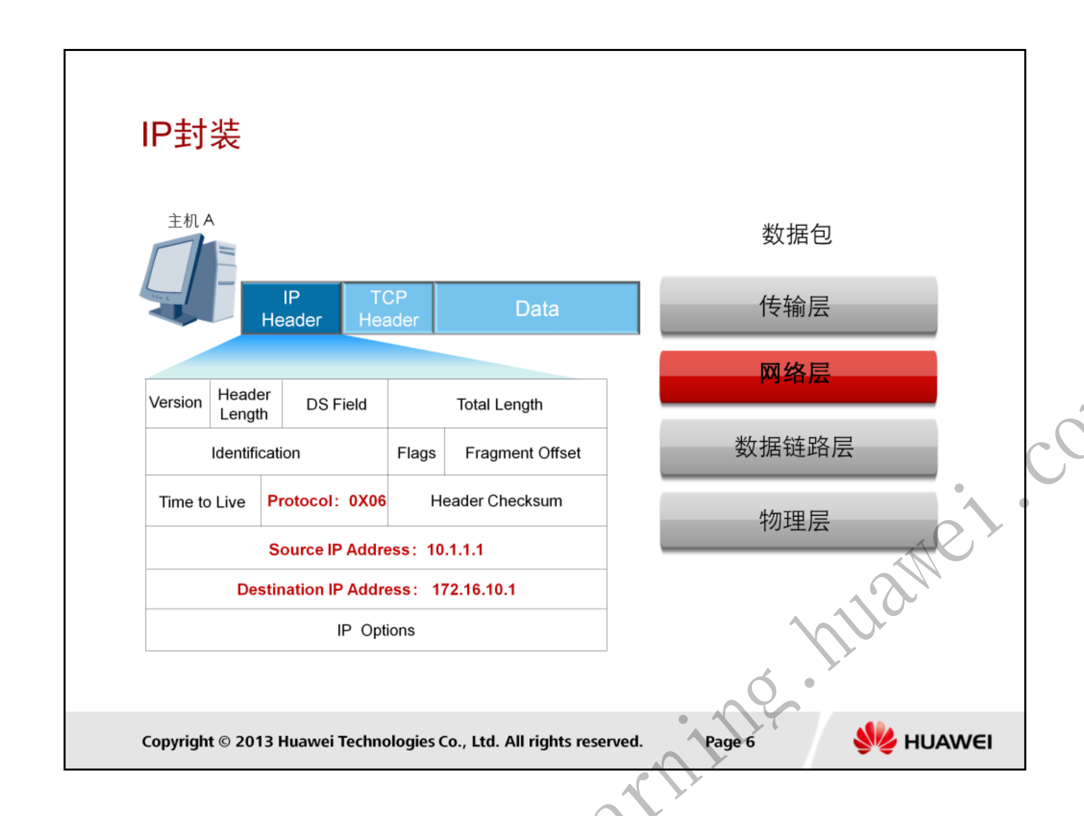


数据可以在同一网络内或者不同网络间传输，数据转发过程也分为本地转发和远程转发，但两者的数据转发原理是基本一样的，都是遵循TCP/IP协议簇。

本示例中，主机A需要访问服务器A的Web服务，并且假定两者之间已经建立了TCP连接。接下来会以此示例来讲解数据在不同网络间的传输过程。



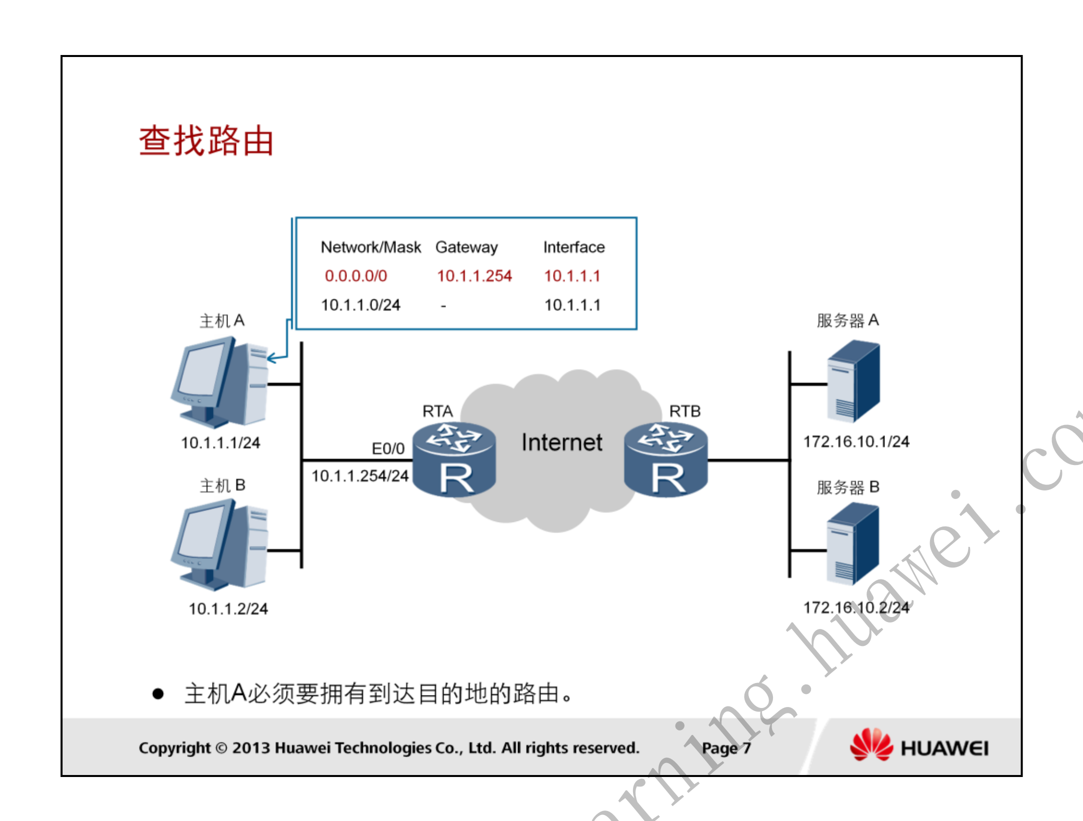
主机A会对待发送的应用数据首先执行加密和压缩等相关操作，之后进行传输层封装。Web应用是基于传输层的TCP协议传输数据的。主机A使用TCP进行报文封装时，必须填充源端口和目的端口字段，初始序列号和确认序列号字段，标识位，窗口字段以及校验和字段。此例中数据段的源端口号为主机A随机选择的1027号端口，目的端口号为服务器A的TCP知名端口80。



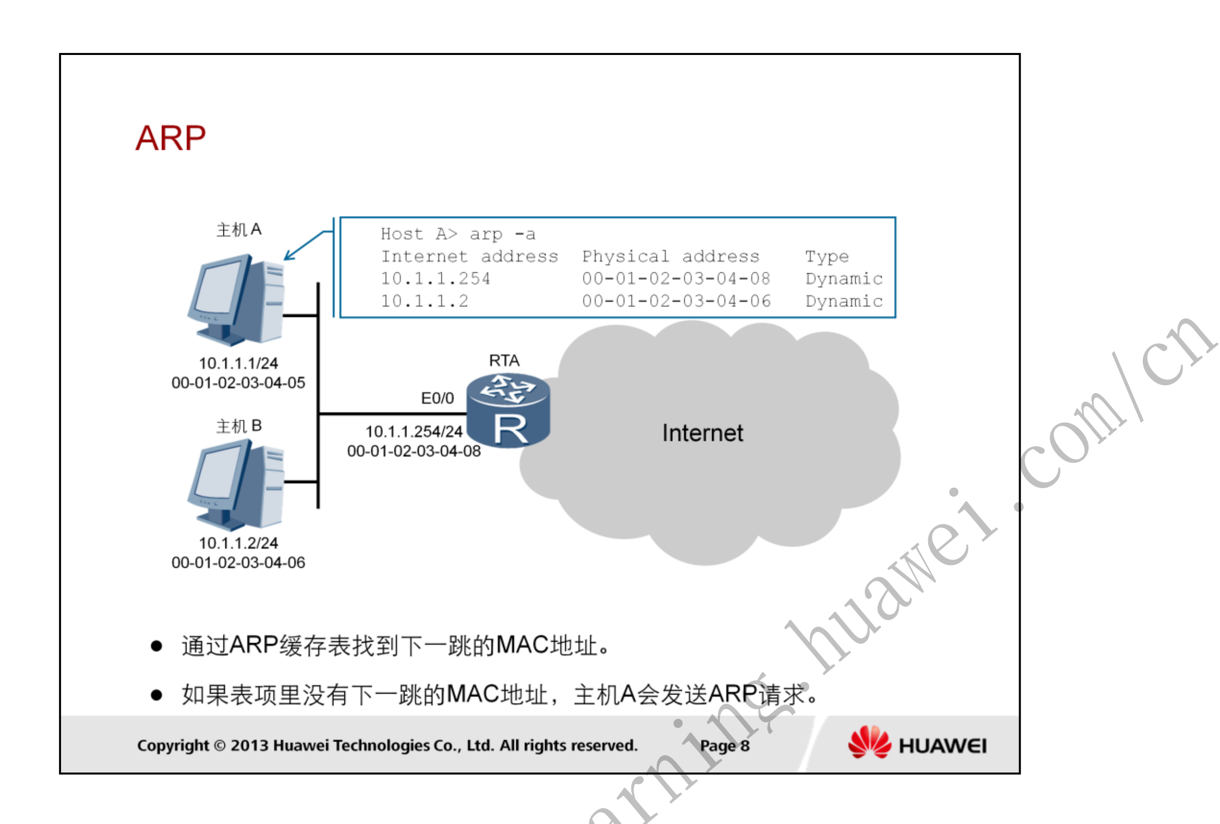
主机A完成传输层封装后，一般会进行网络层数据封装，在使用IP进行封装时，需要明确IP报文的源和目的地址。如果IP报文的大小大于网络的最大传输单元（MTU），则该报文有可能在传输过程中被分片。

生存时间（TTL）字段用来减少网络环路造成的影响。ARG3系列路由器产生的数据包，默认TTL值为255。路由器转发一个数据包时，该值会被减1，如果路由器发现该值被减为0，就会丢弃该数据包。这样，即使网络中存在环路，数据包也不会在网络上一直被转发。

协议字段标识了传输层所使用的协议。本例中，传输层使用的是TCP协议，所以该字段的填充值为0X06。

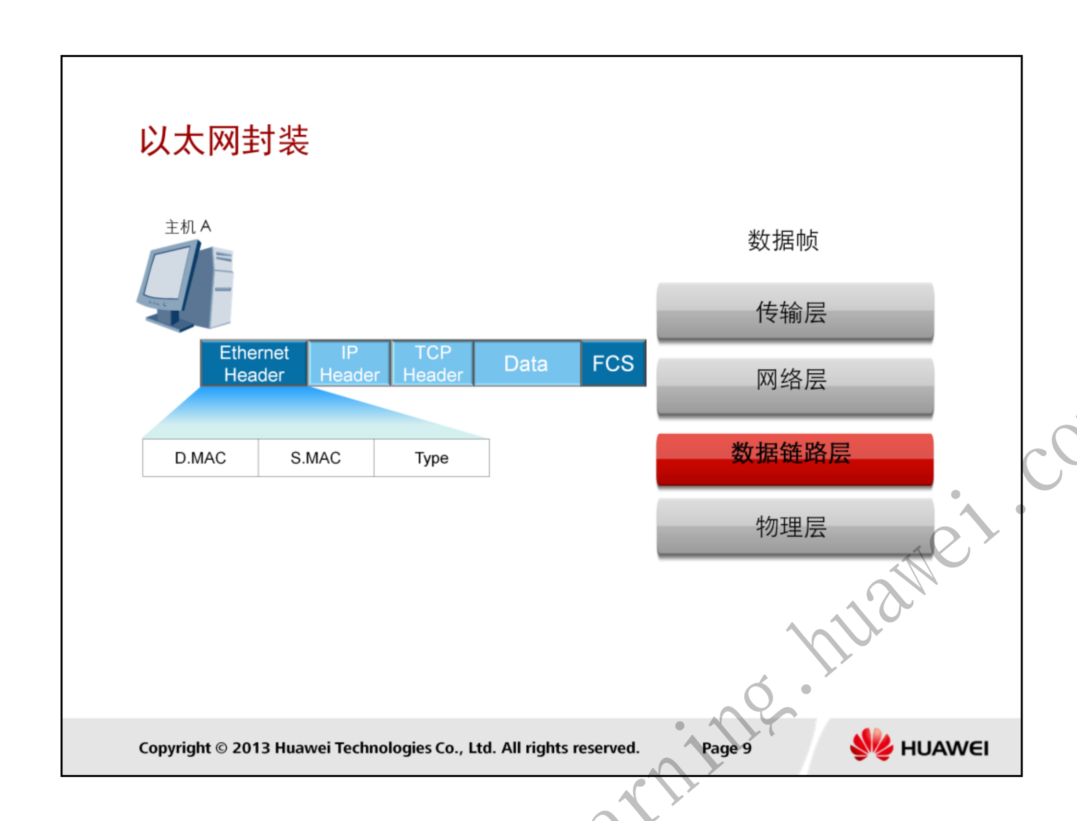


每个主机都会独自维护各自的路由表项。主机A在发送数据前需要先检查是否能够到达目的端，这个过程是通过查找路由来完成的。在此示例中，主机A拥有一条到达“任何网络”（在IP编址部分已经简要介绍过）的路由，它发往其他网络的数据都会通过IP地址为10.1.1.1的接口转发到下一跳，即网关10.1.1.254。



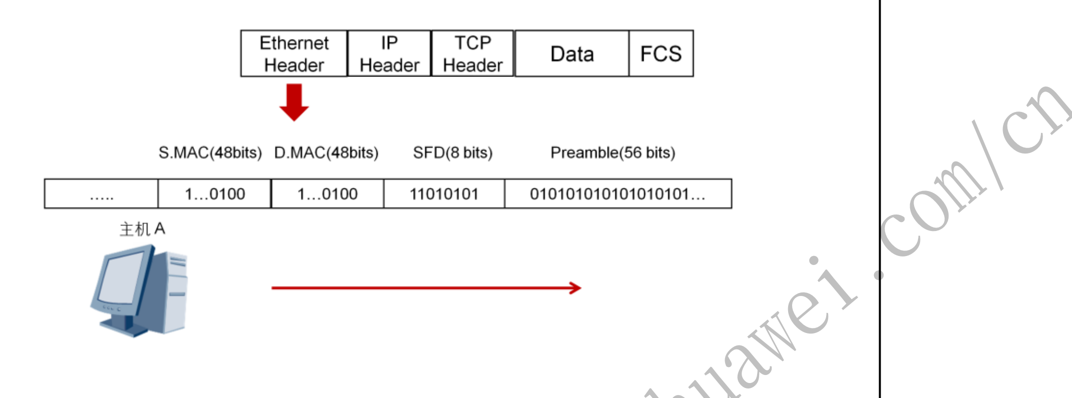
接下来，由于数据包要被封装成数据帧，所以主机A需要获取下一跳的MAC地址，也就是网关的MAC地址。主机首先会查询ARP缓存表。本例中，主机A的ARP缓存表中存在网关MAC地址的表项。

如果没有查找到网关的MAC地址表项，主机A会通过发送ARP请求来获取网关的MAC地址。



主机A在链路层封装数据帧时，会遵循IEEE 802.3或Ethernet_II标准，Ethernet_II帧头中的类型字段填充为0x0800，以表示网络层使用的是IP协议。源MAC地址为主机A的MAC地址，目的MAC地址为网关路由器E0/0接口的MAC地址。

数据帧转发过程

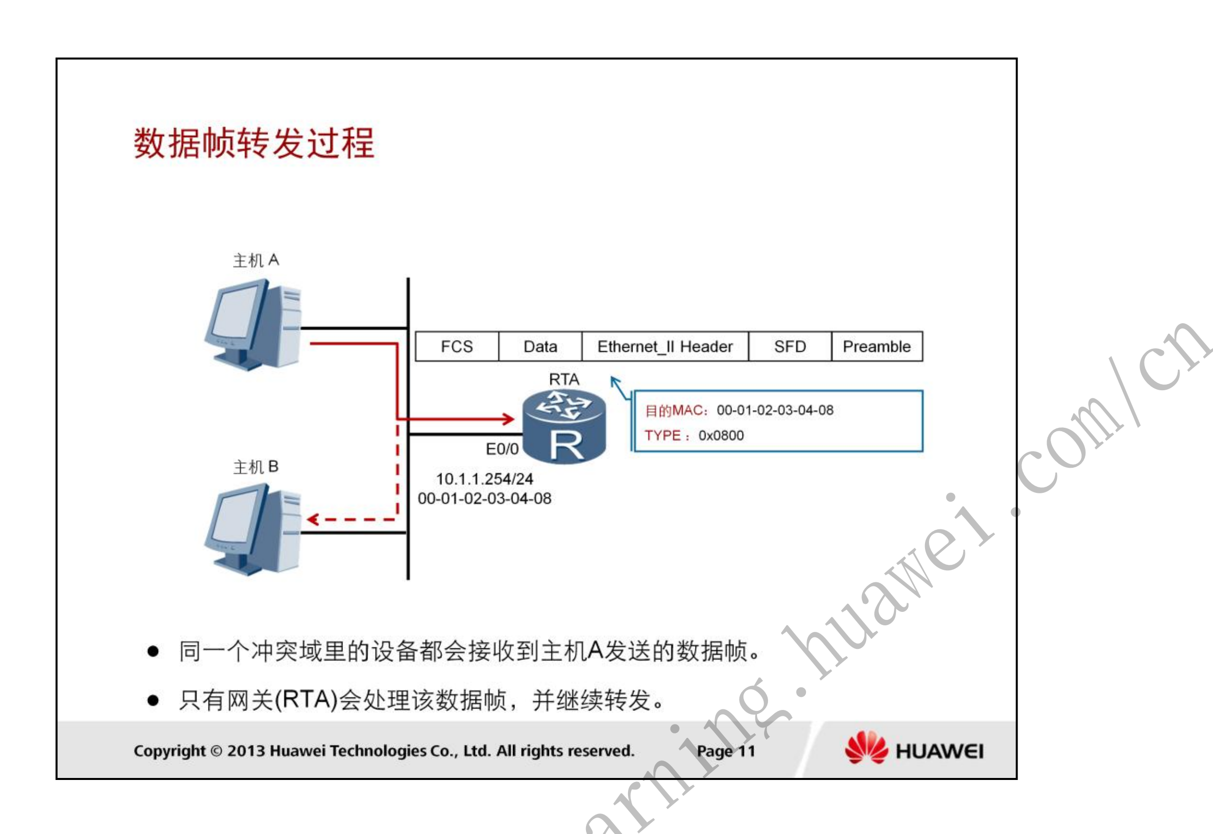


- 主机工作在半双工状态下，suoyi 会使用CSMA/CD来检测链路是否空闲。
- 前导码用于使接收者进入同步状态，定界符用于指示帧的开始。

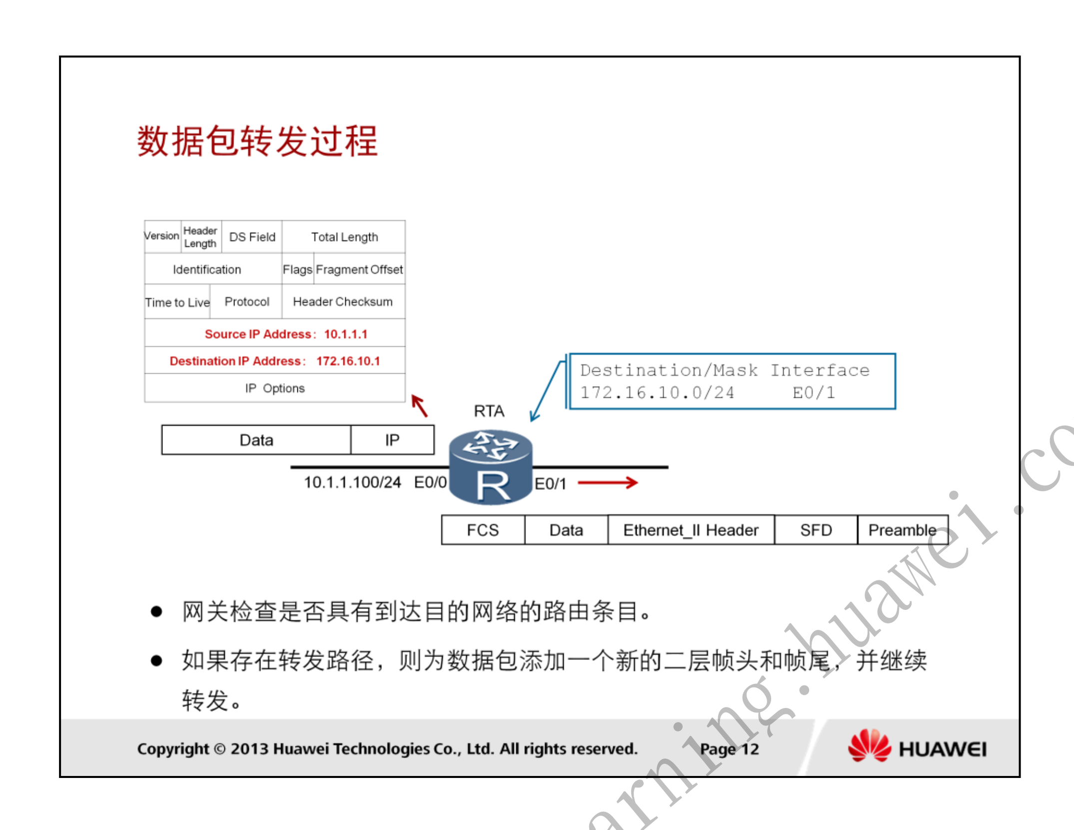
 Page 10



主机A工作在半双工状态下，所以会使用CSMA/CD来检测链路是否空闲。如果链路空闲，主机A会将一个前导码(Preamble)和一个帧首定界符（SFD）附加到帧头然后进行传输。前导码的作用是使接收设备进行同步并做好接收数据帧的准备。前导码是包括了7个字节的二进制“1”、“0”交替的代码，即1010…10共56位。帧首定界符是长度为1个字节的10101011二进制序列，它的作用是使接收端对帧的第一位进行定位。

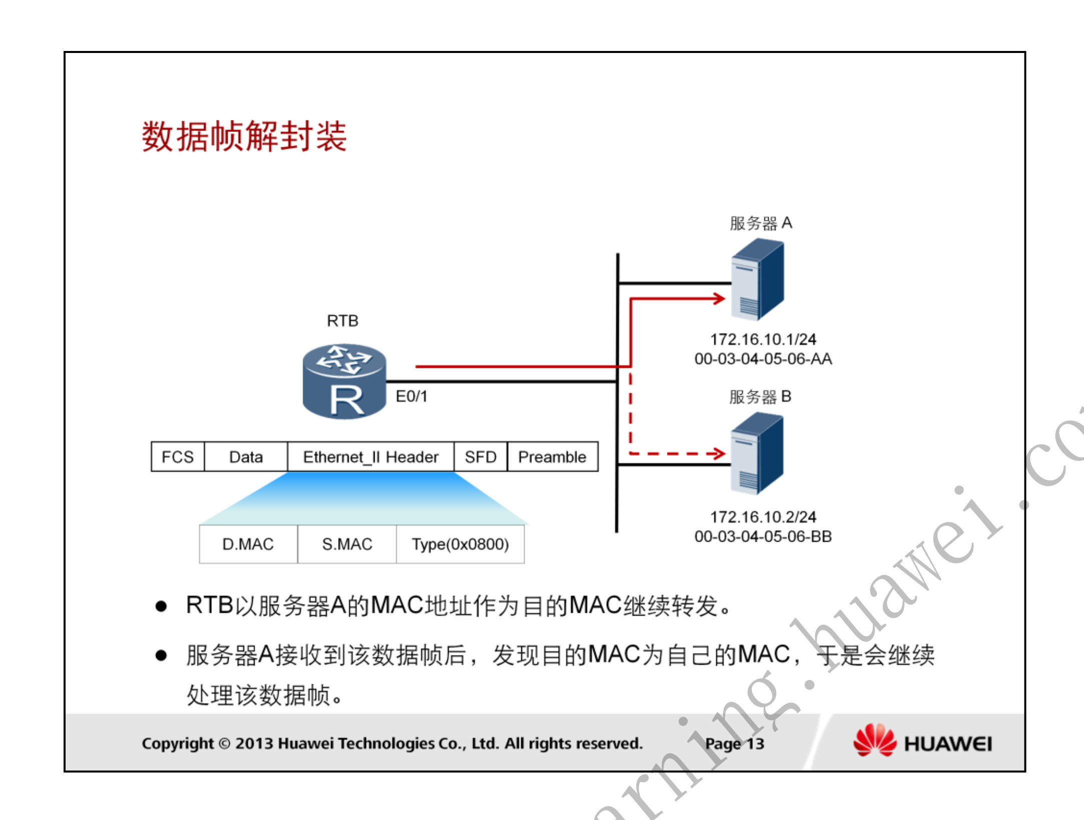


本例中，主机A发送数据帧到共享以太网，此网络中的所有网络设备都会收到该帧。设备收到帧之后，首先会进行FCS校验。如果FCS校验未能通过，则帧被立即丢弃。对于通过了FCS校验的帧，设备会检查帧中的目的MAC地址。如果帧中的目的MAC地址与自己的MAC地址不同，设备将丢弃帧，如果相同，则会继续处理。处理过程中，帧头帧尾会被剥去（也就是解封装），剩下的数据报文会被根据帧头中的类型字段的值来送到网络层中的对应协议模块去处理。



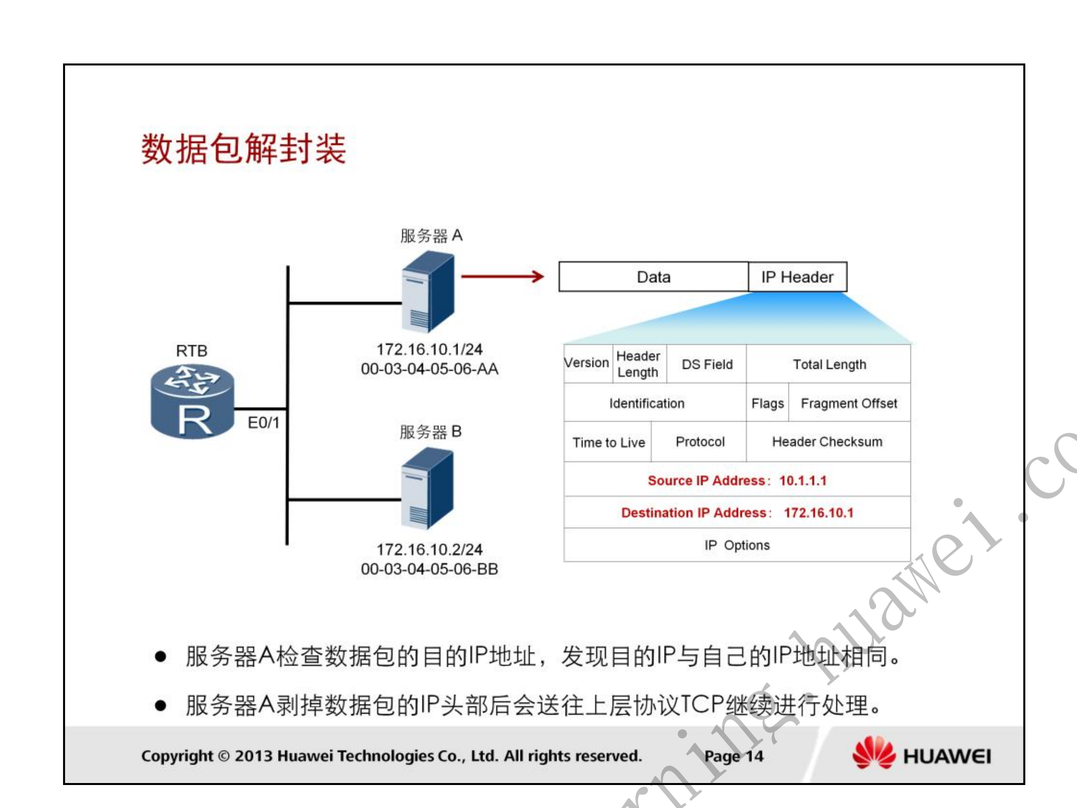
RTA收到此数据报文后，网络层会对该报文进行处理。RTA首先根据IP头部信息中的校验和字段，检查IP数据报文头部的完整性，然后根据目的IP地址查看路由表，确定是否能够将数据包转发到目的端。RTA还必须对TTL的值进行处理。另外，报文大小不能超过MTU值。如果报文大小超过MTU值，则报文将被分片。

网络层处理完成后，报文将被送到数据链路层重新进行封装，成为一个新的数据帧，该帧的头部会封装新的源MAC地址和目的MAC地址。如果当前网络设备不知道下一跳的MAC地址，将会使用ARP来获得。



该示例中，服务器A处于一个共享以太网中，两台服务器都会收到RTB发送的数据帧。该帧的目的MAC地址与服务器B的接口MAC地址不匹配，所以会被服务器B丢弃。

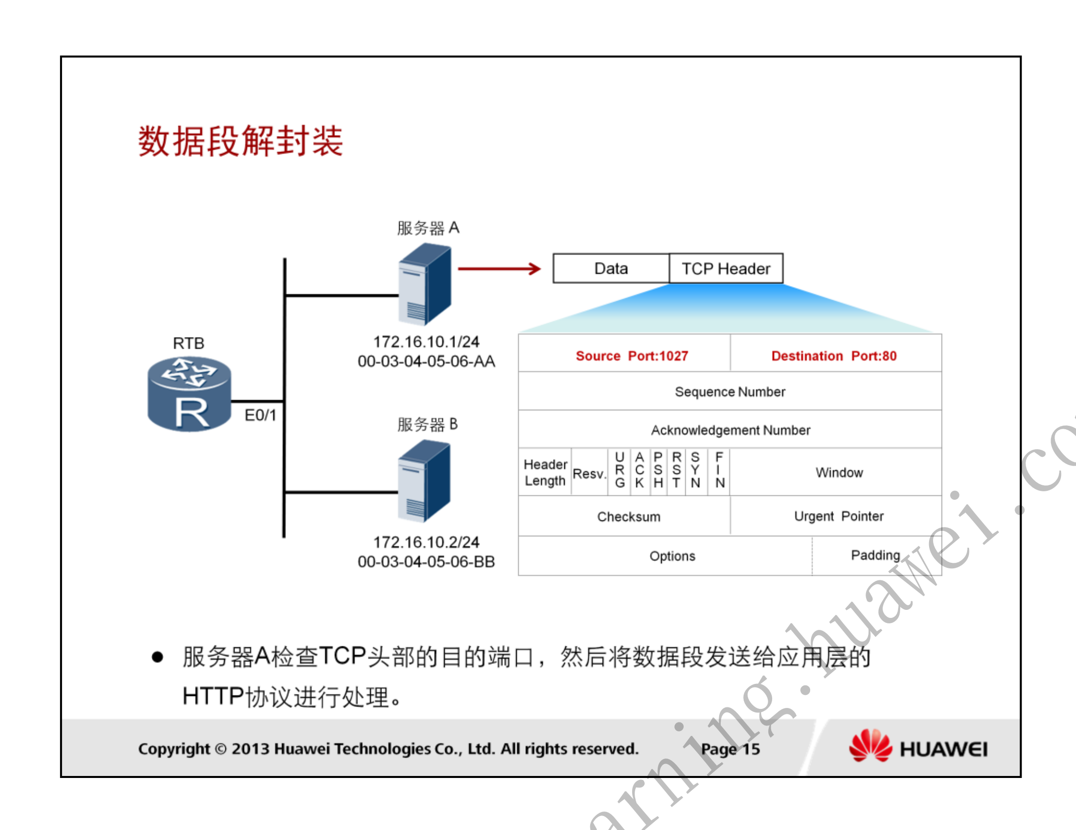
服务器A成功收到该帧，并通过FCS校验。服务器A将利用帧中的类型字段来识别在网络层处理该数据的协议。该示例中，服务器A会将解封装后的此数据交给网络层的IP协议来进行处理。



服务器A通过IP协议来处理该报文，首先会通过校验和字段来验证报文头的完整性，然后检查IP报文头中的目的IP地址是否与自己当前的IP地址匹配。

如果在源与目的之间的数据传输期间数据发生了报文分片，则报文会被目的端重新组合。标识字段用于标识属于同一数据源的分片报文，偏移量表示该分片在原分组中的相对位置。标志字段目前只有两位有意义，标志字段最低位为1表示后面还有分片，为0表示这已经是最后一个数据片；中间一位为1表示不能分片，为0表示允许分片。所有的分片报文必须被目的端全部接收到后才会进行重新组合。

协议字段表示此数据包携带的上层数据是哪种协议的数据。需要注意的是，下一个报头并非总是传输层报头。例如，ICMP报文也是使用IP协议封装，协议字段值为0x01。



当IP报文头被处理完并剥离后，数据段会被发送到传输层进行处理。在此示例中，传输层协议使用的是TCP，且发送端和接收端已经通过三次握手建立了连接。传输层收到该数据段后，TCP协议会查看并处理该数据段头部信息，其中目的端口号为80，用于表示处理该数据的应用层协议为HTTP协议。TCP处理完头部信息后会将此数据段头部进行剥离，然后将剩下的应用数据发送到HTTP协议进行处理。



总结

- 数据在进行二层和三层封装之前，主机需要了解哪些信息?
- 当数据帧发送到非目的主机时，非目的主机将会如何处理?
- 传输层如何能够准确的将数据交给特定应用?
- 当两台主机同时访问服务器的HTTP服务，该服务器如何区分数据属于哪个会话?

 Page 16



1. 主机在封装数据包之前，必须要知道目的端IP地址。在封装数据帧之前，必须要知道去往目的网络的路由以及下一跳的MAC地址。
2. 如果主机接收到一个不是发往自己的数据帧，在检验帧头中的目的MAC地址之后会丢弃该帧。
3. 传输层会检查TCP或UDP报文头中的目的端口号，以此来识别特定应用。
4. 服务器可以只通过源IP地址识别两台主机的HTTP流量，另外TCP报文头中包含的源端口也可以被用来区分同一台主机通过不同的浏览器发起的不同的会话。例如，两个来自源IP为10.1.1.1的HTTP流量使用的目的端口号都是80，但源端口号为1028和1035。


谢谢
www.huawei.com

Module-2

VRP的操作指导


VRP基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

交换机可以隔离冲突域，路由器可以隔离广播域，这两种设备在企业网络中应用越来越广泛。随着越来越多的终端接入到网络中，网络设备的负担也越来越重，这时网络设备可以通过华为专有的VRP系统来提升运行效率。

通用路由平台VRP（Versatile Routing Platform）是华为公司数据通信产品的通用操作系统平台，它以IP业务为核心，采用组件化的体系结构，在实现丰富功能特性的同时，还提供了基于应用的可裁剪和可扩展的功能，使得路由器和交换机的运行效率大大增加。能对VRP熟练地进行配置和操作是对网络工程师的一种基本要求。

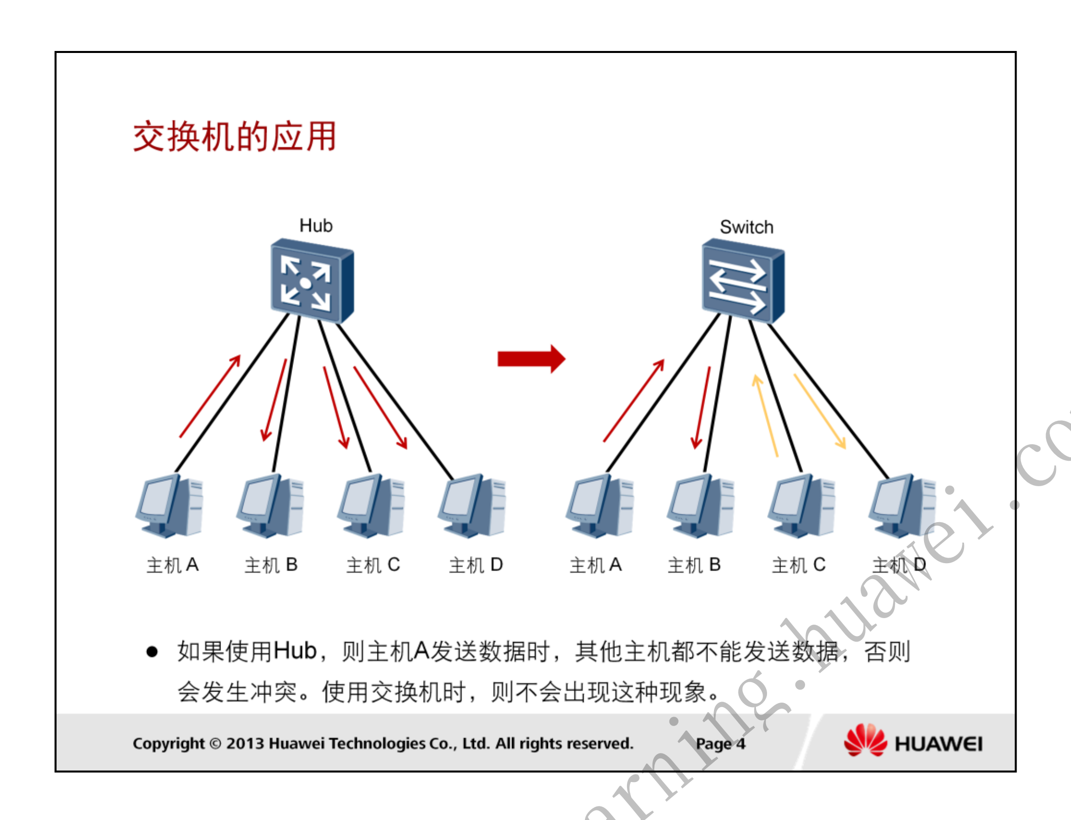




学习目标

学完本课程后，您应该能：

- 掌握交换机和路由器的应用场景
- 掌握冲突域和广播域的区别
- 了解VRP的基础知识

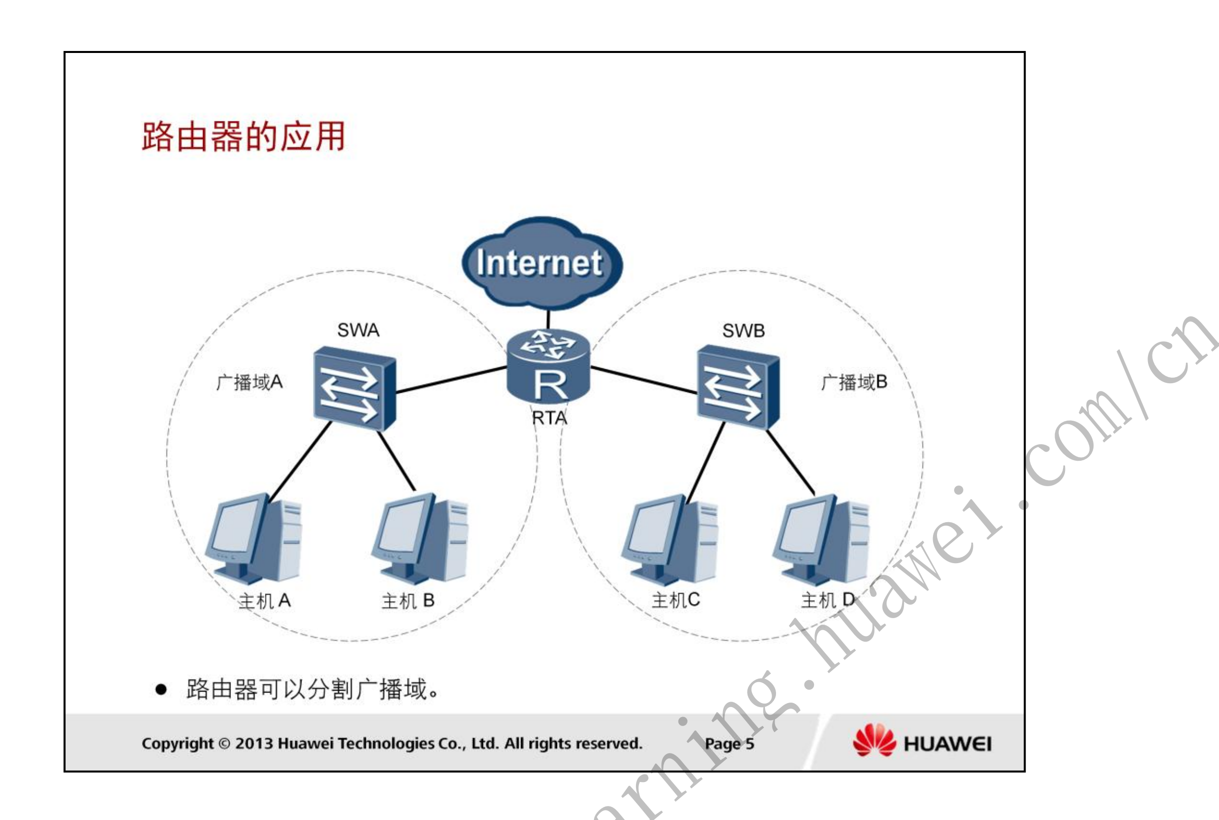


由集线器（HUB）和中继器组建的以太网，实质上是一种共享式以太网。共享式以太网的主要缺陷有：冲突严重、广播泛滥、安全性差。

交换机是工作在数据链路层的设备。交换机可以将一个共享式以太网分割为多个冲突域。链路层流量被隔离在不同的冲突域中进行转发，如此便极大地提升了以太网的性能。

更进一步说，通常主机和交换机之间以及交换机与交换机之间都使用全双工技术进行通信，这时冲突现象会被彻底消除。

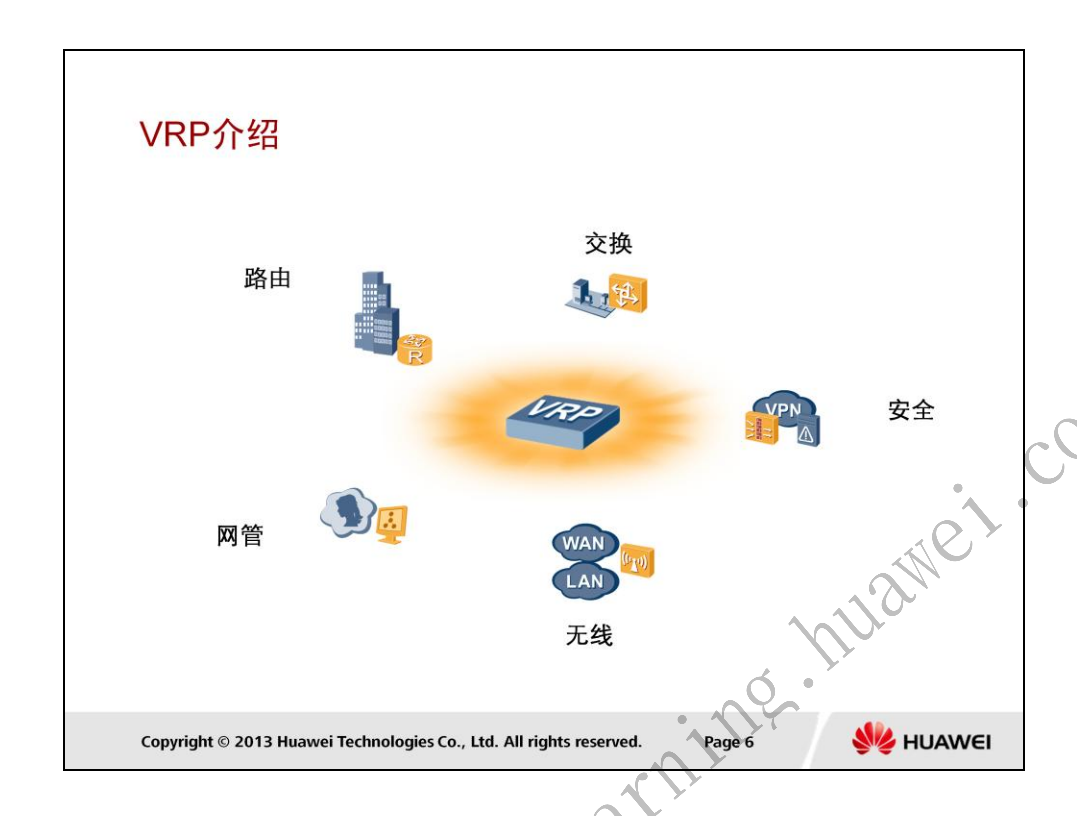
如本例所示，在由Hub搭建的网络中，所有的主机处于同一个冲突域，主机A发送数据给主机B时，其他主机都将收到此数据，但同时这些主机都不能发送数据。用交换机替代Hub后，因为交换机分割了冲突域，所以在主机A发送数据给主机B时，主机C和主机D之间也可以同时互相发送数据。



交换机虽然能够隔离冲突域，但是当一台设备发送广播帧时，其他设备仍然都会接收到该广播帧。随着网络规模的增大，广播会越来越多，这样就会影响网络的效率。路由器可以用来分割广播域，减少广播对网络效率的影响。

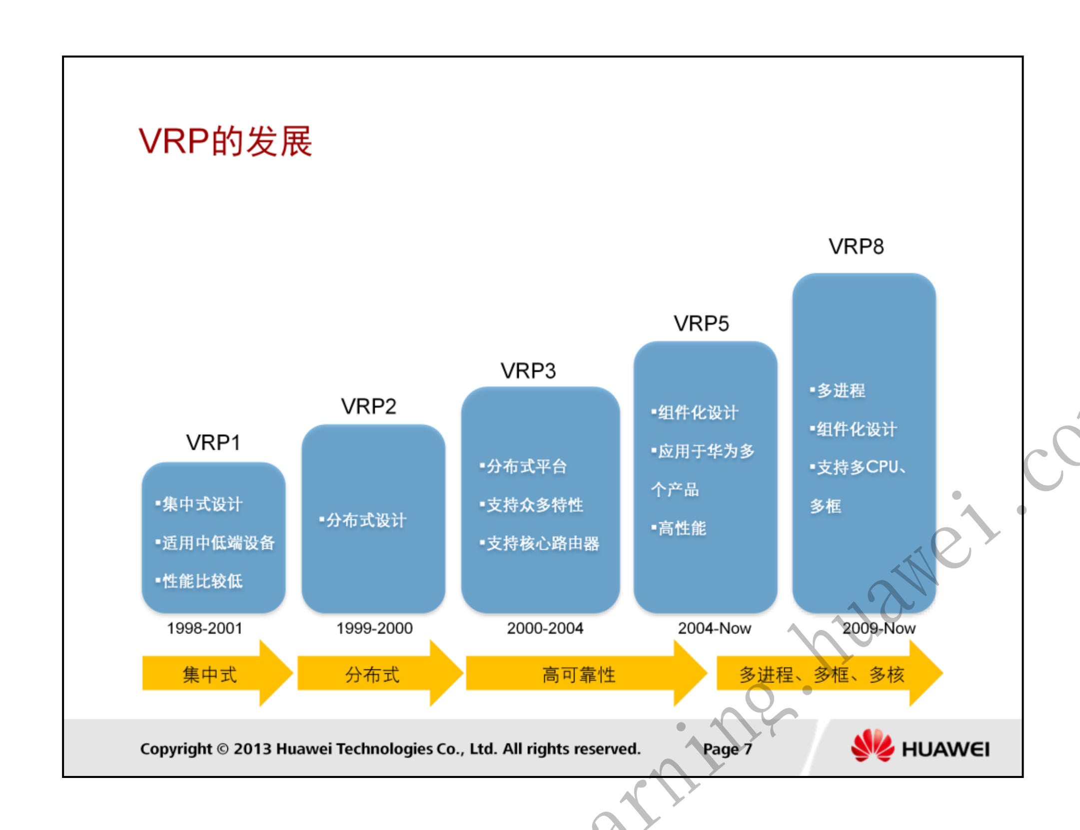
一般情况下，广播帧的转发被限制在广播域内。广播域的边缘是路由器，因为通常路由器不会转发广播帧。

路由器负责在网络间转发报文。它能够在自身的路由表里查找到达目的地的下一跳地址，将报文转发给下一跳路由器，如此重复，并最终将报文送达目的地。



VRP是华为公司具有完全自主知识产权的网络操作系统，可以运行在多种硬件平台之上。VRP拥有一致的网络界面、用户界面和管理界面，为用户提供了灵活丰富的应用解决方案。

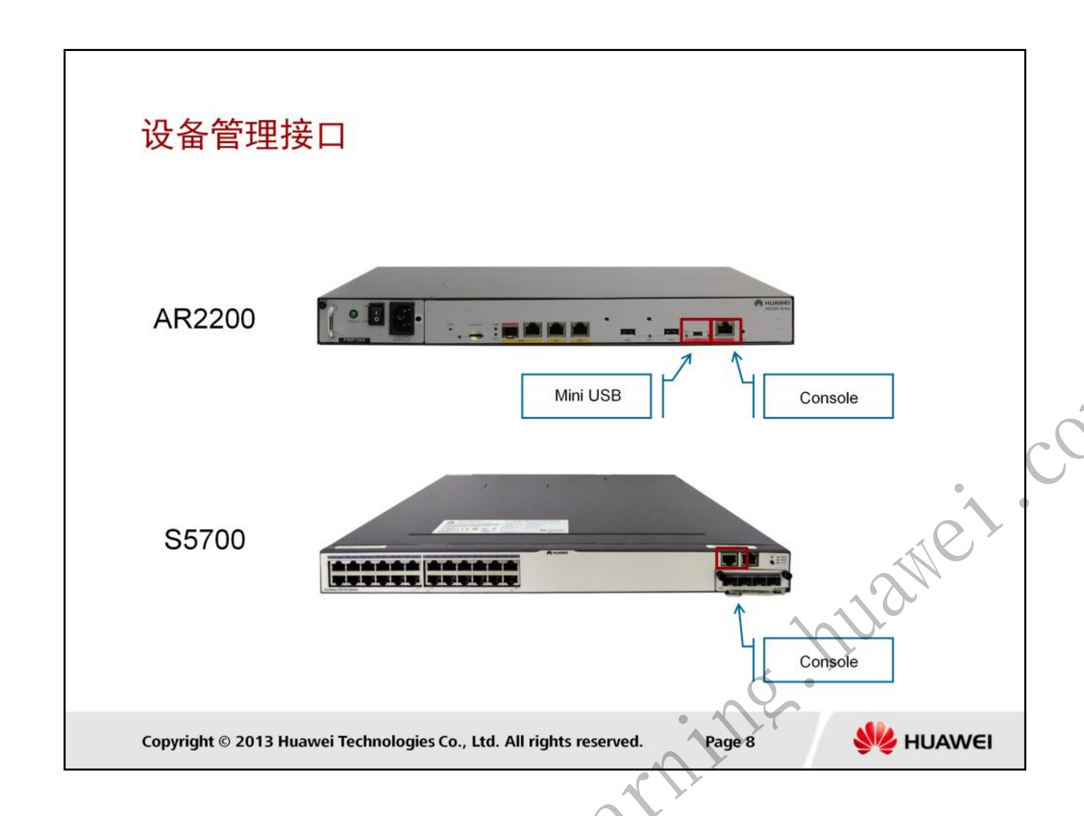
VRP平台以TCP/IP协议簇为核心，实现了数据链路层、网络层和应用层的多种协议，在操作系统中集成了路由交换技术、QoS技术、安全技术和IP语音技术等数据通信功能，并以IP转发引擎技术作为基础，为网络设备提供了出色的数据转发能力。



随着网络技术和应用的飞速发展，VRP平台在处理机制、业务能力、产品支持等方面也在持续演进。到目前为止，VRP已经开发出了5个版本，分别是VRP1、VRP2、VRP3、VRP5和VRP8。

VRP5是一款分布式网络操作系统，具有高可靠性、高性能、可扩展的架构设计。目前，绝大多数华为设备使用的都是VRP5版本。

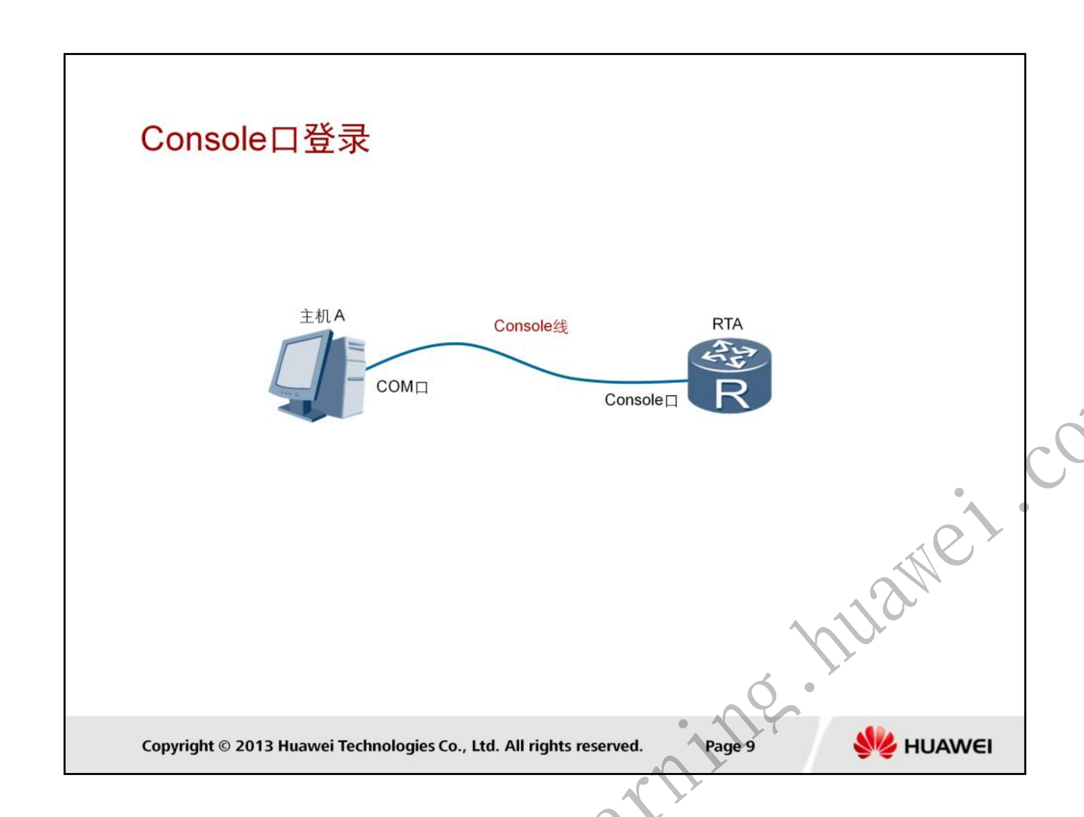
VRP8是新一代网络操作系统，具有分布式、多进程、组件化架构，支持分布式应用和虚拟化技术，能够适应未来的硬件发展趋势和企业急剧膨胀的业务需求。



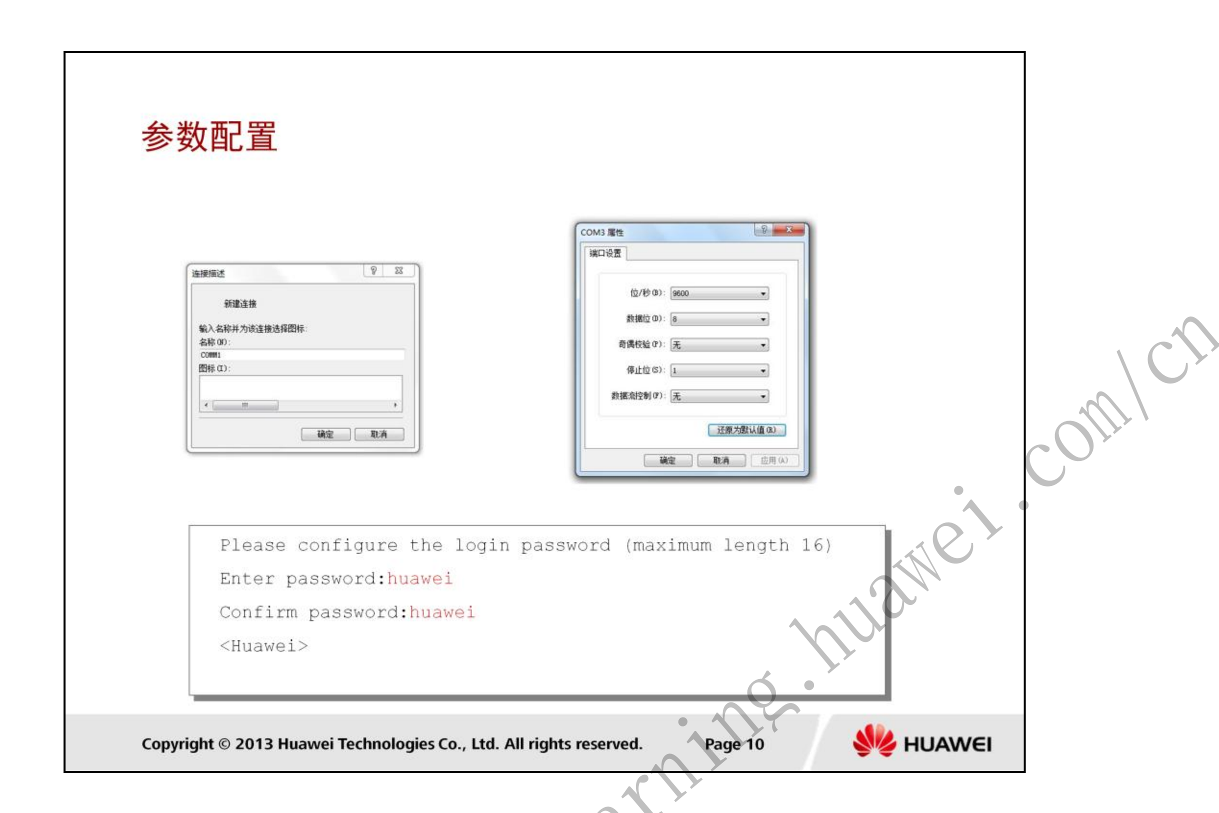
AR系列企业路由器有多个型号，包括AR150、AR200、AR1200、AR2200、AR3200。它们是华为第三代路由器产品，提供路由、交换、无线、语音和安全等功能。AR路由器被部署在企业网络和公网之间，作为两个网络间传输数据的入口和出口。在AR路由器上部署多种业务能降低企业的网络建设成本和运维成本。根据一个企业的用户数和业务的复杂程度可以选择不同型号的AR路由器来部署到网络中。

华为X7系列以太网交换机提供数据交换的功能，满足企业网络上多业务的可靠接入和高质量传输的需求。这个系列的交换机定位于企业网络的接入层、汇聚层和核心层，提供大容量交换，高密度端口，实现高效的报文转发。X7系列以太网交换机包括了S1700、S2700、S3700、S5700、S7700、S9700等。

ARG3系列路由器和X7系列交换机都提供了Console口作为管理口，AR2200额外提供了Mini USB口作为管理口。



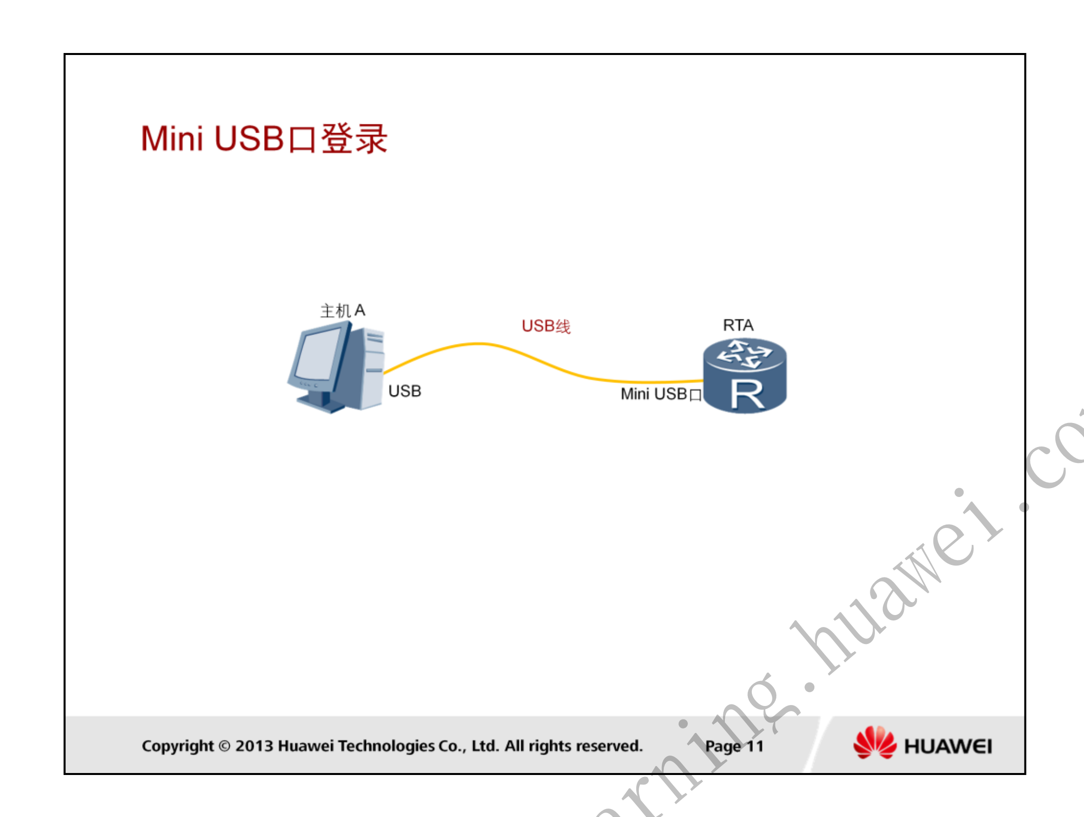
使用Console线缆来连接交换机或路由器的Console口与计算机的COM口，这样就可以通过计算机实现本地调试和维护。S5700和AR2200的Console口是一种符合RS232串口标准的RJ45接口。目前大多数台式电脑提供的COM口都可以与Console口连接。笔记本电脑一般不提供COM口，需要使用USB到RS232的转换接口。



很多终端模拟程序都能发起Console连接，例如，可以使用超级终端程序连接到VRP操作系统，如上图所示。使用超级终端连接VRP时，必须设置端口参数。上图是端口参数设置的示例，如果对参数值做了修改，需要恢复默认参数值。

完成设置以后，点击“确定”按钮即可与VRP建立连接。如果设备初次启动，VRP系统会要求用户设置Console登录密码。

在缺少超级终端程序的计算机上，可以使用putty或Secure CRT程序发起Console连接，并连接到VRP，配置参数与上图一致。



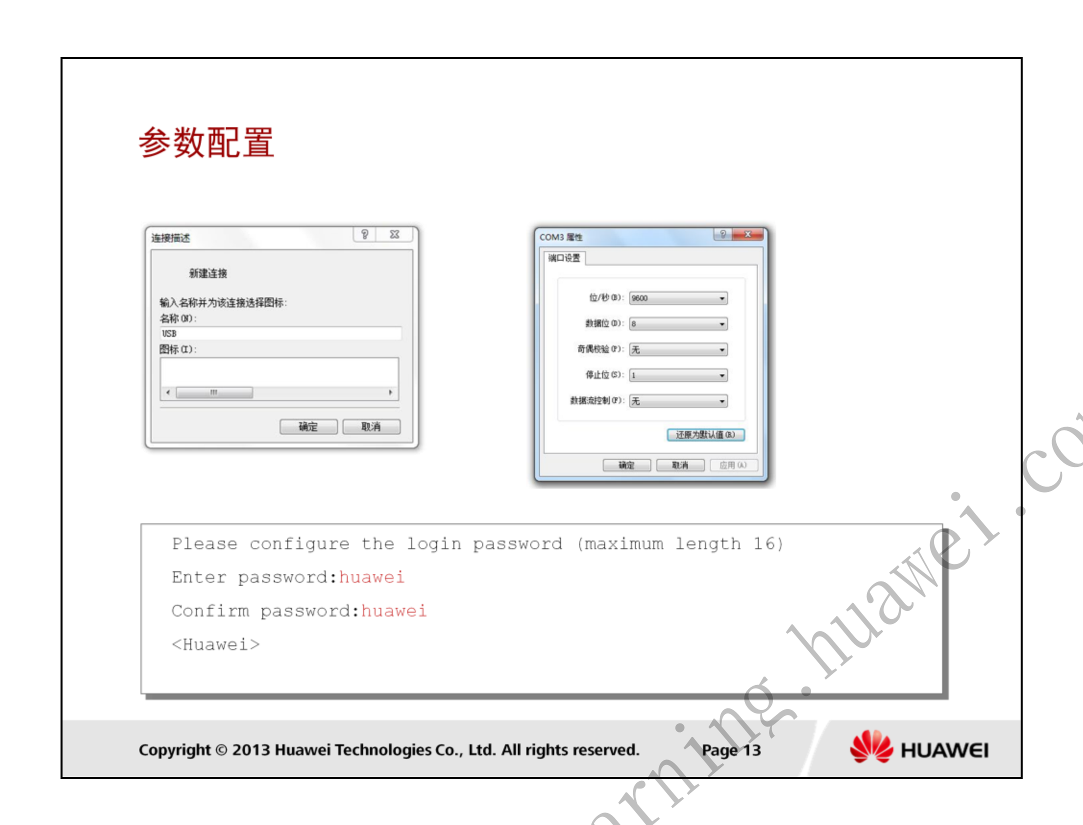
华为AR2200系列路由器还支持通过Mini USB口与主机USB口建立连接，实现对设备的调试和维护。

在管理设备时，Console接口和Mini USB接口互斥，即同一时刻只能使用其中的1个接口连接到VRP。



在使用Mini USB口建立连接前，需要在主机上安装驱动程序。您可以从华为企业官方支持网站下载到所需驱动程序。目前，Mini USB的驱动程序只能安装在Windows XP、Windows Vista和Windows 7操作系统上。按照软件提示安装驱动程序即可。

华为企业官方支持网站网址：http://www.support.huawei.com/enterprise/。



安装驱动程序后，主机上会增加一个新的虚拟COM接口，终端模拟软件可以通过该虚拟COM接口连接到VRP。具体的软件使用和参数配置与本材料的第10页一致。



总结

- 如果路由器收到了网络中主机发送的广播报文，会如何操作?
- 华为数通设备目前使用的VRP版本是多少?



1. 当路由器收到该广播报文时，路由器会根据数据包内容进行处理，可能会对必要广播报文（如请求路由器MAC地址的ARP广播）进行回应，但不会将该数据包转发到其他广播域。
2. 目前，大多数华为数通产品使用的是VRP5版本，少数产品如NE系列路由器使用的是VRP8版本。


谢谢
www.huawei.com


命令行基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

熟悉VRP命令行并且熟练掌握VRP配置是高效管理华为网络设备的必备基础。

HUAWEI



学习目标

学完本课程后，您应该能：

- 掌握VRP命令行的基础知识
- 利用VRP命令行进行基本的配置

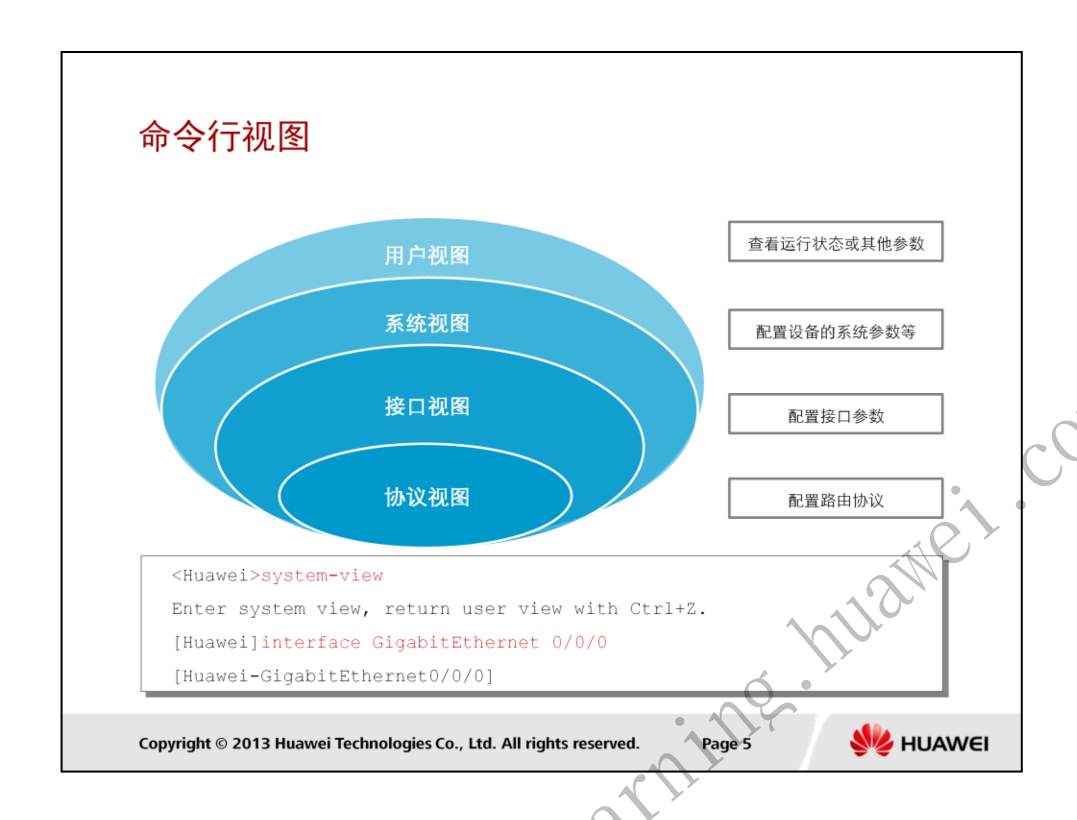
设备初始化启动

```
BIOS Creation Date : Jan  5 2013, 18:00:24
DDR DRAM init : OK
Start Memory Test ? ('t' or 'T' is test):skip
Copying Data : Done
Uncompressing : Done
……
Press Ctrl+B to break auto startup ... 1
Now boot from sd1:/AR2220-V200R003C00SPC200.cc,
……
<Huawei>
 Warning: Auto-Config is working. Before configuring the device, stop
Auto-Config. If you perform configurations when Auto-Config is
running, the DHCP, routing, DNS, and VTY configurations will be lost.
Do you want to stop Auto-Config? [y/n]:Y
```

 Page 4



管理员和工程师如果要访问在通用路由平台VRP上运行的华为产品，首先要进入启动程序。开机界面信息提供了系统启动的运行程序和正在运行的VRP版本及其加载路径。启动完成以后，系统提示目前正在运行的是自动配置模式。用户可以选择是继续使用自动配置模式还是进入手动配置的模式。如果选择手动配置模式，在提示符处输入Y。在没有特别要求的情况下，我们选择手动配置模式。



VRP分层的命令结构定义了很多命令行视图，每条命令只能在特定的视图中执行。本例介绍了常见的命令行视图。每个命令都注册在一个或多个命令视图下，用户只有先进入这个命令所在的视图，才能运行相应的命令。进入到VRP系统的配置界面后，VRP上最先出现的视图是用户视图。在该视图下，用户可以查看设备的运行状态和统计信息。

若要修改系统参数，用户必须进入系统视图。用户还可以通过系统视图进入其他的功能配置视图，如接口视图和协议视图。

通过提示符可以判断当前所处的视图，例如：“< >”表示用户视图，“[ ]”表示除用户视图以外的其它视图。

命令行功能

命令	功能
CTRL+A	把光标移动到当前命令行的最前端
CTRL+C	停止当前命令的运行
CTRL+Z	回到用户视图
CTRL+]	终止当前连接或切换连接

```
<Huawei>system-view
Enter system view, return user view with Ctrl+Z.
[Huawei]^z  //Ctrl+Z
<Huawei>
```

 Page 6 

为了简化操作，系统提供了快捷键，使用户能够快速执行操作。以上表格中提供了系统定义的快捷键。其他的快捷键功能如下：

CTRL+B 将光标向左移动一个字符。

CTRL+D 删除当前光标所在位置的字符。

CTRL+E 将光标移动到当前行的末尾。

CTRL+F 将光标向右移动一个字符。

CTRL+H 删除光标左侧的一个字符。

CTRL+N 显示历史命令缓冲区中的后一条命令。

CTRL+P 显示历史命令缓冲区中的前一条命令。

CTRL+W 删除光标左侧的一个字符串。

CTRL+X 删除光标左侧所有的字符。

CTRL+Y 删除光标所在位置及其右侧所有的字符。

ESC+B 将光标向左移动一个字符串。

ESC+D 删除光标右侧的一个字符串。

ESC+F 将光标向右移动一个字符串。

命令行功能

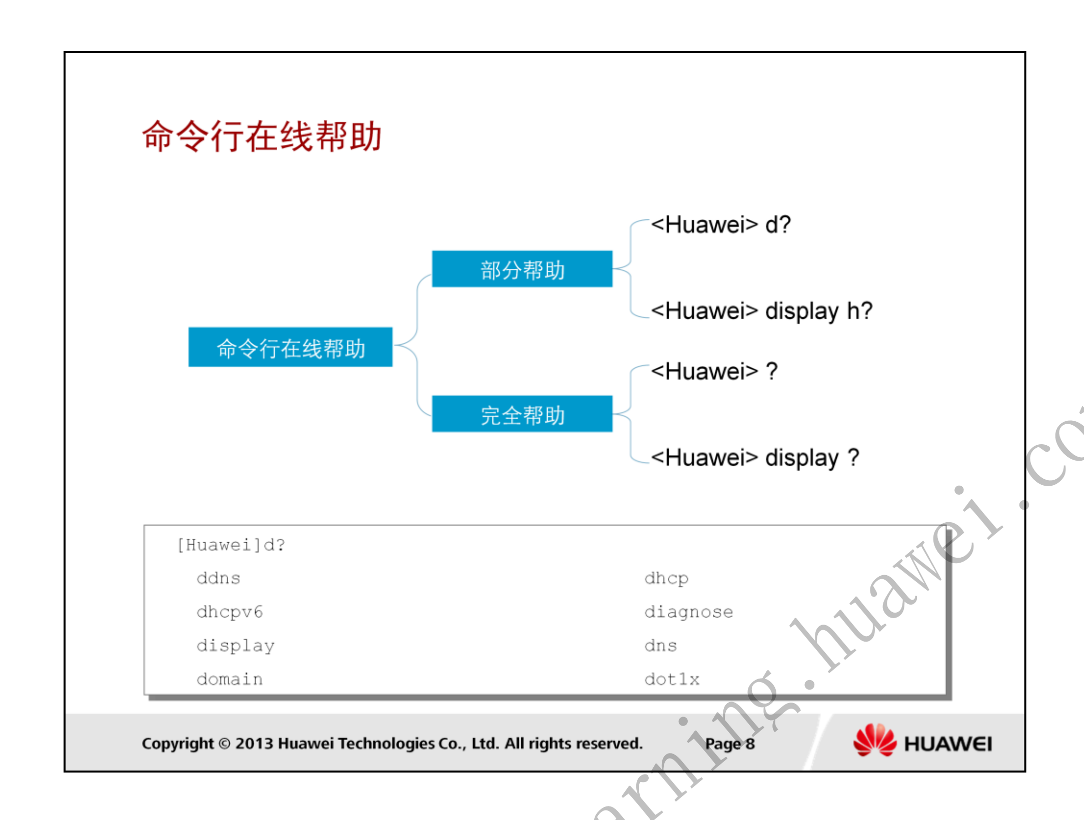
命令	功能
Backspace	删除光标左边的第一个字符
← or Ctrl+B	光标左移一位
→ or Ctrl+F	光标右移一位
TAB	输入一个不完整的命令并按TAB键，就可以补全该命令

```
[Huawei]inter    //TAB
[Huawei]interface
```




还有一些其他功能键也可以用来执行类似的操作，比如，与CTRL+H的功能一样，退格也可以删除光标左侧的一个字符。向左的光标键(←)与向右的光标键(→)可以分别用来执行与CTRL+B和CTRL+F相同的功能。向下的光标键(↓)可以用来执行与Ctrl+N相同的功能，向上的光标键(↑)可以替换CTRL+P。

此外，若命令字的前几个字母是独一无二的，系统可以在输完该命令的前几个字母后自动将命令补充完整。如本例所示，用户只需输入inter并按Tab键，系统自动将命令补充为interface。若命令字并非独一无二的，按Tab键后将显示所有可能的命令。如输入in并按Tab键，系统会按顺序显示以下命令： info-center，interface。



VRP提供两种帮助功能，分别是部分帮助和完全帮助。

部分帮助指的是，当用户输入命令时，如果只记得此命令关键字的开头一个或几个字符，可以使用命令行的部分帮助获取以该字符串开头的所有关键字的提示，如本例中所示。

完全帮助指的是，在任一命令视图下，用户可以键入“?”获取该命令视图下所有的命令及其简单描述；如果键入一条命令关键字，后接以空格分隔的“?”，如果该位置为关键字，则列出全部关键字及其描述。



网络上一般都会部署不止一台设备，管理员需要对这些设备进行统一管理。在进行设备调试的时候，首要任务是设置设备名。设备名用来唯一地标识一台设备。AR2200路由器默认的设备名是huawei，而S5700系列默认的设备名是Quidway。设备名称一旦设置，立刻生效。

配置系统时钟

命令	功能
clock timezone	设置所在时区
clock datetime	设置当前时间和日期
clock daylight-saving-time	设置采用夏时制

```
<Huawei>clock timezone BJ add 08:00:00
<Huawei>clock datetime 10:20:29 2013-04-11
<Huawei>display clock
2013-04-11 10:20:48
Thursday
Time Zone(BJ) : UTC+08:00
```

 Page 10 

系统时钟是设备上的系统时间戳。由于地域的不同，用户可以根据当地规定设置系统时钟。用户必须正确设置系统时钟以确保其与其他设备保持同步。

设置系统时钟的公式为：UTC+时区偏移量+夏时制时间偏移量。**clock datetime**命令设置*HH:MM:SS YYYY-MM-DD*格式的系统时钟。但是需要注意的是，如果没有设定时区，或者时区设定为零，那么设定的日期和时间将被认为是UTC时间，所以建议在对系统时间和日期进行配置前先设置时区。

clock timezone命令用来对本地时区信息进行设置，具体的命令参数为*time-zone-name* { *add* | *minus* } *offset*。其中参数add表示与UTC时间相比，*time-zone-name*增加的时间偏移量。即，在系统默认的UTC时区的基础上，加上*offset*，就可以得到*time-zone-name*所标识的时区时间；参数minus指的是与UTC时间相比，*time-zone-name*减少的时间偏移量。即，在系统默认的UTC时区的基础上，减去*offset*，就可以得到*time-zone-name*所标识的时区时间。

有的地区实行夏令时制，因此当进入夏令时实施区间的一刻，系统时间要根据用户的设定进行夏令时时间的调整。VRP支持夏令时功能。比如，在英国，从三月的最后一个星期天到十月最后一个星期天是夏令时区间，那么可以通过执行命令指定夏令时的开始和结束时间。

配置标题消息

命令	功能
header login	配置在用户登陆前显示的标题消息
header shell	配置在用户登陆后显示的标题消息

```
[Huawei]header login information "welcome to huawei certification!"
[Huawei]header shell information "Please don't reboot the device!"
……
welcome to huawei certification!
Login authentication
Password:
Please don't reboot the device!
<Huawei>
```

 Page 11 

header命令用来设置用户登录设备时终端上显示的标题信息。

login参数指定当用户在登录设备认证过程中，激活终端连接时显示的标题信息。

shell参数指定当用户成功登录到设备上，已经建立了会话时显示的标题信息。

header的内容可以是字符串或文件名。当header的内容为字符串时，标题信息以第一个英文字符作为起始符号，最后一个相同的英文字符作为结束符；通常情况下，建议使用英文特殊符号，并需要确保在信息正文中没有此符号。

本例中，header的内容是字符串。字符串可以包含1-2000字符，包含空格。使用header { login | shell } information *text*命令能设置字符串形式的header。

若要设置文件形式的header，使用header { *login* | *shell* } file *file-name*命令。*file-name*参数指定了标题信息所使用的文件名，登陆前后，该文件的内容将以文本的形式显示出来。

命令等级

用户等级	命令等级	名称
0	0	访问级
1	0 and 1	监控级
2	0,1 and 2	配置级
3-15	0,1,2 and 3	管理级

```
<Huawei> system-view
[Huawei]command-privilege level 3 view user save
```

 Page 12 

系统将命令进行分级管理，以增加设备的安全性。设备管理员可以设置用户级别，一定级别的用户可以使用对应级别的命令行。缺省情况下命令级别分为0～3级，用户级别分为0～15级。用户0级为访问级别，对应网络诊断工具命令（ping、tracert）、从本设备出发访问外部设备的命令（Telnet客户端）、部分display命令等。用户1级为监控级别，对应命令级0、1级，包括用于系统维护的命令以及display等命令。用户2级是配置级别，包括向用户提供直接网络服务，包括路由、各个网络层次的命令。用户3-15级是管理级别，对应命令3级，该级别主要是用于系统运行的命令，对业务提供支撑作用，包括文件系统、FTP、TFTP下载、文件交换配置、电源供应控制，备份板控制、用户管理、命令级别设置、系统内部参数设置、以及用于业务故障诊断的debugging命令。本例展示了如何修改命令级别，在用户视图下执行save命令需要3级的权限。

在具体使用中，如果我们有多个管理员帐号，但只允许某一个管理员保存系统配置，则可以将save命令的级别提高到4级，并定义只有该管理员有4级权限。这样，在不影响其他用户的情况下，可以实现对命令的使用控制。

用户界面

用户界面类型	编号
Console	0
VTY	0-4

```
<Huawei>system-view
[Huawei]user-interface vty 0 4
[Huawei-ui-vty0-4]
```

- VTY 接口最大可配范围为0-14。

 Page 13 

每类用户界面都有对应的用户界面视图。用户界面（User-interface）视图是系统提供的一种命令行视图，用来配置和管理所有工作在异步交互方式下的物理接口和逻辑接口，从而达到统一管理各种用户界面的目的。在连接到设备前，用户要设置用户界面参数。系统支持的用户界面包括Console用户界面和VTY用户界面。控制口（Console Port）是一种通信串行端口，由设备的主控板提供。虚拟类型终端（Virtual Type Terminal）是一种虚拟线路端口，用户通过终端与设备建立Telnet或SSH连接后，也就建立了一条VTY，即用户可以通过VTY方式登录设备。设备一般最多支持15个用户同时通过VTY方式访问。执行**user-interface maximum-vty** *number* 命令可以配置同时登录到设备的VTY类型用户界面的最大个数。如果将最大登录用户数设为0，则任何用户都不能通过Telnet或者SSH登录到路由器。*display user-interface* 命令用来查看用户界面信息。

不同的设备，或使用不同版本的VRP软件系统，具体可以被使用的VTY接口的最大数量可能不同。

配置用户界面命令

命令	功能
idle-timeout	设置超时时间
screen-length	设置指定终端屏幕的临时显示行数
history-command max-size	设置历史命令缓冲区的大小

```
# Set the size of the history command buffer to 20.
<Huawei>system-view
[Huawei]user-interface console 0
[Huawei-ui-console0]history-command max-size 20
# Set the timeout duration to 1 minute and 30 seconds.
[Huawei-ui-console0]idle-timeout 1 30
```

 Page 14 

用户可以设置Console界面和VTY界面的属性，以提高系统安全性。如果一个连接上设备的用户一直处于空闲状态而不断开，可能会给系统带来很大风险，所以在等待一个超时时间后，系统会自动中断连接。这个闲置切断时间又称超时时间，默认为10分钟。

当display命令输出的信息超过一页时，系统会对输出内容进行分页，使用空格键切换下一页。

如果一页输出的信息过少或过多时，用户可以执行screen-length命令修改信息输出时一页的行数。默认行数为24，最大支持512行。不建议将行数设置为0，因为那样将不会显示任何输出内容了。

每条命令执行过后，执行的记录都保存在历史命令缓存区。用户可以利用(↑)，(↓)　，CTRL+P，Ctrl+N这些快捷键调用这些命令。历史命令缓存区中默认能存储10条命令，可以通过运行*history-command max-size*改变可存储的命令数，最多可存储256条。

配置登陆权限

命令	功能
user privilege	配置指定用户界面下的用户级别
set authentication password	配置本地认证密码

```
# Set the user level on the VTY0 user interface to 2.
<Huawei>system-view
[Huawei]user-interface vty 0
[Huawei-ui-vty0]user privilege level 2
[Huawei-ui-vty0-4]set authentication password cipher huawei
```

 Page 15



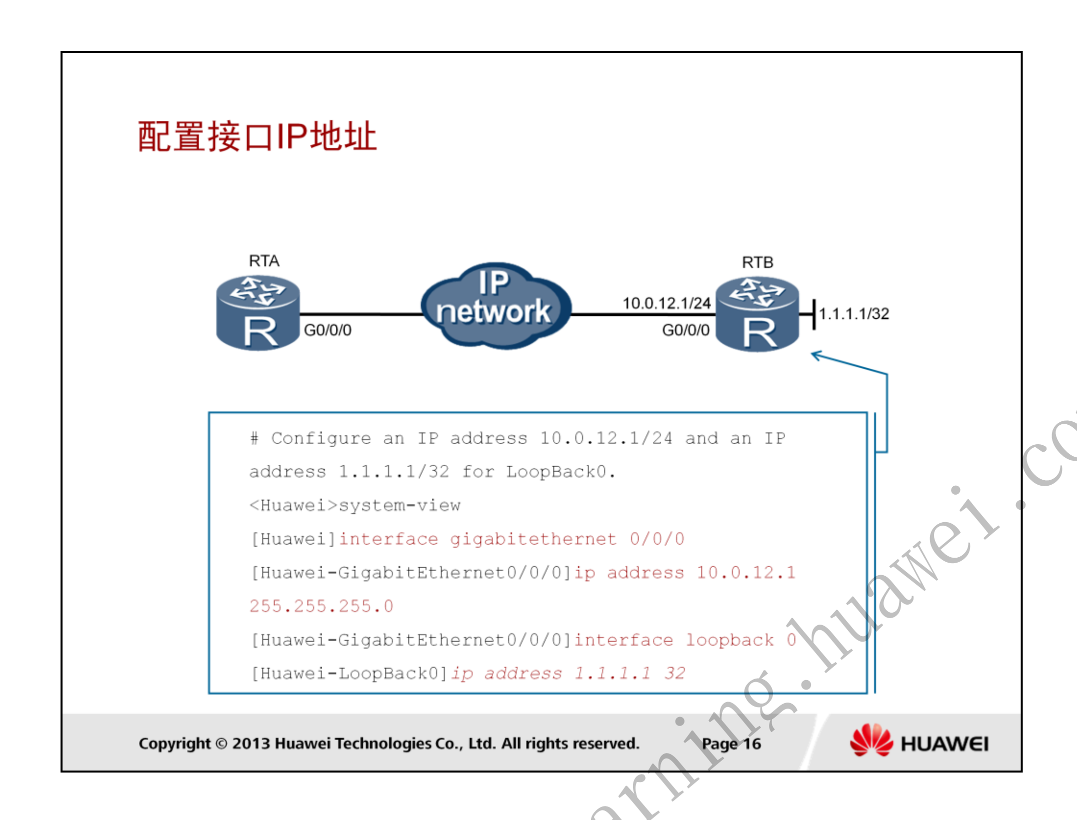
本页介绍只使用密码登陆的情况下，登陆权限的密码配置方式。

如果没有权限限制，未授权的用户就可以使用设备获取信息并更改配置。从设备安全的角度考虑，限制用户的访问和操作权限是很有必要的。用户权限和用户认证是提升终端安全的两种方式。用户权限要求规定用户的级别，一定级别的用户只能执行特定级别的命令。

配置用户界面的用户认证方式后，用户登录设备时，需要输入密码进行认证，这样就限制了用户访问设备的权限。在通过VTY进行Telnet连接时，所有接入设备的用户都必须要经过认证。

设备提供三种认证模式，AAA模式、密码认证模式和不认证模式。AAA认证模式具有很高的安全性，因为登录时必须输入用户名和密码。密码认证只需要输入登录密码即可，所以所有的用户使用的都是同一个密码。使用不认证模式就是不需要对用户认证直接登陆到设备。需要注意的是，Console界面默认使用不认证模式。

对于Telnet登录用户，授权是非常必要的，最好设置用户名、密码和指定和帐号相关联的权限。



要在接口运行IP服务，必须为接口配置一个IP地址。一个接口一般只需要一个IP地址。在特殊情况下，也有可能为接口配置一个次要IP地址。例如，当路由器AR2200的接口连接到一个物理网络时，该物理网络中的主机属于两个网段。为了让两个网段的主机都可以通过路由器AR2200访问其它网络，可以配置一个主IP地址和一个次要IP地址。一个接口只能有一个主IP地址，如果接口配置了新的主IP地址，那么新的主IP地址就替代了原来的主IP地址。

用户可以利用ip address <*ip-address* > { mask | *mask-length* } 命令为接口配置IP地址，这个命令中，*mask*代表的是32 比特的子网掩码，如255.255.255.0，*mask-length* 代表的是可替换的掩码长度值，如24，这两者可以交换使用。

Loopback接口是一个逻辑接口，可用来虚拟一个网络或者一个IP主机。在运行多种协议的时候，由于Loopback接口稳定可靠，所以也可以用来做管理接口。

在给物理接口配置IP地址时，需要关注该接口的物理状态。默认情况下，华为路由器和交换机的接口状态为up；如果该接口曾被手动关闭，则在配置完IP地址后，应使用undo shutdown打开该接口。



总结

- 华为网络设备支持多少个用户同时使用console口登录?
- 在使用命令interface loopback interface 0 之后，loopback 0接口的状态是什么?

 Page 17



1. 华为网络设备同时只能有一个用户登录Console界面，因此Console用户的编号固定为0。
2. Loopback接口是一种逻辑接口，在未创建之前，Loopback接口并不存在。从创建开始，Loopback接口就一直存在，并一直保持Up状态，除非被手动关闭。


谢谢
www.huawei.com



文件系统基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

华为网络设备的配置文件和VRP系统文件都保存在物理存储介质中，所以文件系统是VRP正常运行的基础。只有掌握了对文件系统的基本操作，网络工程师才能对设备的配置文件和VRP系统文件进行高效的管理。





学习目标

学完本课程后，您应该能：

- 掌握文件系统的基本操作

基本查询命令

功能	命令
查看当前目录	pwd
显示当前目录下的文件信息	dir
查看文本文件的具体内容	more

```
<Huawei>dir
Directory of flash:/
  Idx    Attr      Size(Byte)  Date           Time        FileName
    0    drw-             -    Apr 10 2013 09:30:35       src
    1    -rw-            28    Apr 10 2013 09:31:38       private-data.txt
    2    -rw-           120    Apr 10 2013 09:32:38       wzbk1.cfg
……
32,004 KB total (31,995 KB free)
```




VRP基于文件系统来管理设备上的文件和目录。在管理文件和目录时，经常会使用一些基本命令来查询文件或者目录的信息，常用的命令包括**pwd**，**dir** [/all] [*filename* | *directory*]和
more [/binary] *filename* [*offset*] [all]。

pwd命令用来显示当前工作目录。

dir [/all] [*filename* | *directory*]命令用来查看当前目录下的文件信息。

more [/binary] *filename* [*offset*] [all]命令用来查看文本文件的具体内容。

本例中，在用户视图中使用dir命令，可以查看flash中的文件信息。

目录操作

功能	命令
修改用户当前界面的工作目录	cd
创建新的目录	mkdir
删除目录	rmdir

```
<Huawei>mkdir test
Info: Create directory flash:/test......Done.
<Huawei>dir
Directory of flash:
  Idx    Attr      Size(Byte)   Date        Time        FileName
    0    drw-          -      Apr 10 2013 09:30:35    src
    1    -rw-         28      Apr 10 2013 09:31:38    private-data.txt
    2    -rw-        120      Apr 10 2013 09:32:38    wzbk1.cfg
    3    drw-          -      Apr 10 2013 09:53:11    test
......
32,004 KB total (31,995 KB free)
```

 Page 5 

目录操作常用的命令包括：**cd** *directory*，**mkdir** *directory*和**rmdir** *directory*。

cd *directory*命令用来修改用户当前的工作目录。

mkdir *directory*命令能够创建一个新的目录。目录名称可以包含1-64个字符。

rmdir *directory*命令能够删除文件系统中的目录，此处需要注意的是，只有空目录才能被删除。

此例中使用**mkdir** test创建了一个新的目录test，通过**dir**可以查看到新目录test已经创建成功。

文件操作

功能	命令
复制文件	copy
移动文件	move
重命名文件	rename

```
<Huawei>rename test.txt huawei.txt
Rename flash:/test.txt to flash:/huawei.txt ?[Y/N]:y
Info: Rename file flash:/test.txt to flash:/huawei.txt ......Done.
<Huawei>dir
Directory of flash:/
  Idx   Attr     Size(Byte)  Date          Time       FileName
    0   drw-              -  Apr 10 2013 09:30:35   src
    1   -rw-             28  Apr 10 2013 09:31:38   private-data.txt
    2   -rw-            120  Apr 10 2013 09:32:38   wzbkl.cfg
    3   -rw-             12  Apr 10 2013 09:53:11   huawei.txt
......
32,004 KB total (31,995 KB free)
```

 Page 6



文件操作包括：复制、移动、重命名、压缩、删除、恢复等。

copy *source-filename destination-filename*命令可以复制文件。如果目标文件已存在，系统会提示此文件将被替换。目标文件名不能与系统启动文件同名，否则系统将会出现错误提示。

move *source-filename destination-filename*命令可以用来将文件移动到其他目录下。**move**命令只适用于在同一储存设备中移动文件。

rename *old-name new-name*命令可以用来对目录或文件进行重命名。

本例中使用了**rename**命令修改test.txt的名称为huawei.txt。

文件操作

功能	命令
删除/永久删除文件	delete /unreserved
恢复删除的文件	undelete
彻底删除回收站中的文件	reset recycle-bin

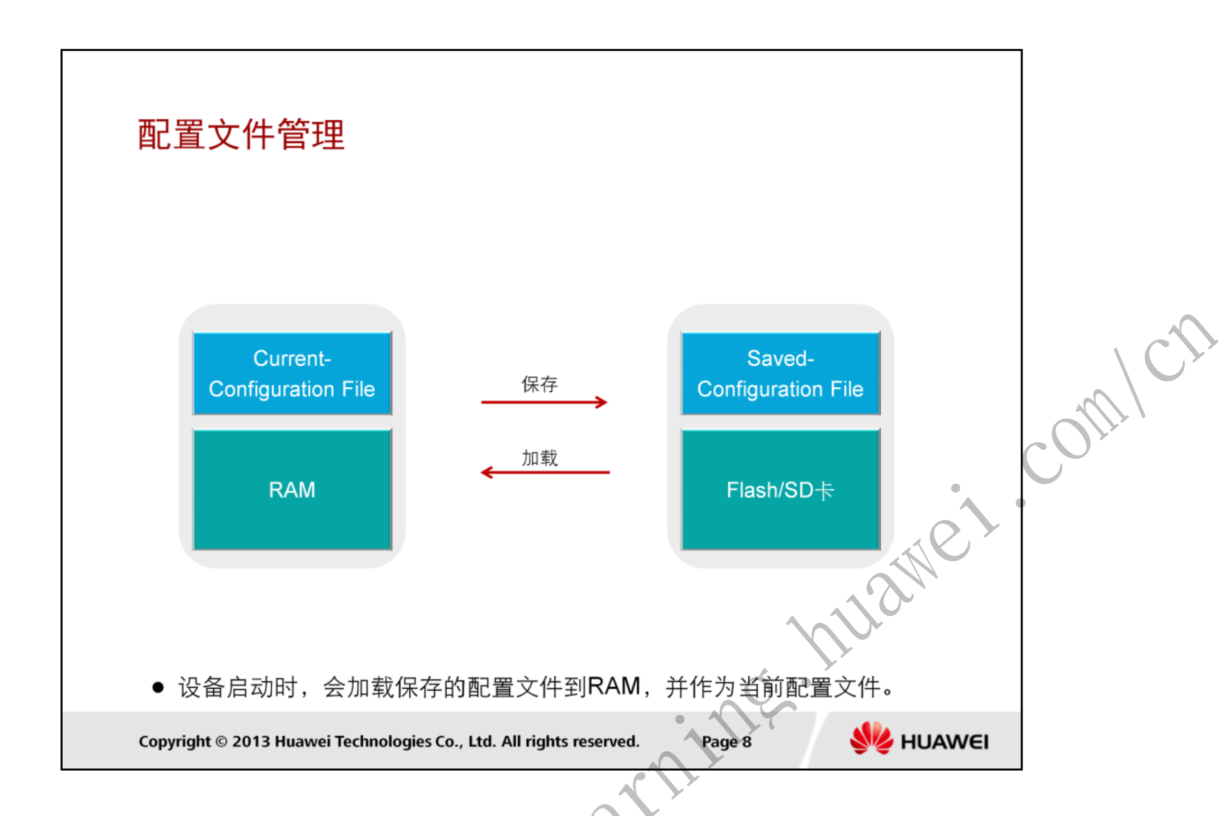
```
<Huawei>delete /unreserved flash:/huawei.txt
<Huawei>dir
Directory of flash:/
  Idx    Attr      Size(Byte)  Date          Time         FileName
    0    drw-            -    Apr 10 2013 09:30:35   src
    1    -rw-           28    Apr 10 2013 09:31:38   private-data.txt
    2    -rw-          120    Apr 10 2013 09:32:38   wzbk1.cfg
……
32,004 KB total (30,995 KB free)
```

 Page 7



delete [/unreserved] [/force] { *filename* | *devicename* }命令可以用来删除文件。一般情况下，被删除的文件将直接被移动到回收站。回收站中的文件也可以通过执行undelete命令进行恢复，但是如果执行delete命令时指定了unreserved参数，则文件将被永久删除。在删除文件时，系统会提示“是否确定删除文件”，但如果命令中指定了/force 参数，系统将不会给出任何提示信息。*filename*参数指的是需要删除的文件的名称，*device-name*参数指定了储存设备的名称。

reset recycle-bin [*filename* | *devicename*]可以用来永久删除回收站中的文件，*filename*参数指定了需要永久删除的文件的名称，*device-name*参数指定了储存设备的名称。



设备中的配置文件分为两种类型：当前配置文件和保存的配置文件。当前配置文件储存在设备的RAM中。用户可以通过命令行对设备进行配置，配置完成后使用save命令保存当前配置到存储设备中，形成保存的配置文件。保存的配置文件都是以“.cfg”或“.zip”作为扩展名，存放在存储设备的根目录下。

在设备启动时，会从默认的存储路径下加载保存的配置文件到RAM中。如果默认存储路径中没有保存的配置文件，则设备会使用缺省参数进行初始化配置。

配置文件查询

功能	命令
显示当前配置文件	display current-configuration
显示保存的配置文件	display saved-configuration

```
<Huawei>display current-configuration
#
sysname Huawei
……
#
Return
<Huawei>display saved-configuration
#
sysname Huawei
……
#
Return
```

 Page 9 

display current-configuration命令可以用来查看设备当前生效的配置。

display current-configuration | begin {*regular-expression*} 命令可以显示以不同参数或表达式开头的配置。

display current-configuration | include {*regular-expression*}命令可以显示包含了指定关键字或表达式的配置。

display saved-configuration [*last*|*time*]命令用来查看设备下次启动时加载的配置文件。使用last参数可以显示本次启动时使用的配置文件内容。使用time参数可以显示系统启动后最近的一次手工或者系统自动保存配置的时间。

配置文件保存

功能	命令
保存当前配置信息	save

```
<Huawei>save
 The current configuration will be written to the device.
  Are you sure to continue? (y/n)[n]:y
  It will take several minutes to save configuration file, please
  wait...............
  Configuration file had been saved successfully
  Note: The configuration file will take effect after being activated
```

 Page 10 

save [*configuration-file*]命令可以用来保存当前配置信息到系统默认的存储路径中。configuration-file为配置文件的文件名，此参数可选。

本例中，执行save命令后，当前配置被保存到了设备的默认储存路径，默认文件名为vrpcfg.zip。



display startup命令用来查看设备本次及下次启动相关的系统软件、备份系统软件、配置文件、License文件、补丁文件以及语音文件。

Startup system software表示的是本次系统启动所使用的VRP文件。

Next startup system software表示的是下次系统启动所使用的VRP文件。

Startup saved-configuration file表示的是本次系统启动所使用的配置文件。

Next startup saved-configuration file表示的是下次系统启动所使用的配置文件。



设备启动时，会从存储设备中加载配置文件并进行初始化。如果存储设备中没有配置文件，设备将会使用默认参数进行初始化。

startup saved-configuration [*configuration-file*] 命令用来指定系统下次启动时使用的配置文件，configuration-file参数为系统启动配置文件的名称。



compare configuration [*configuration-file*] [*current-line-number save-line-number*]命令用来比较当前的配置与下次启动的配置文件内容的区别，*configuration-file*指定需要与当前配置进行比较的配置文件名，*current-line-number*表示从当前配置的该行号开始比较，*save-line-number*表示从指定配置的该行号开始比较。

当执行该命令后，系统默认会将保存的配置与当前配置从第一行开始逐行进行比较。如果指定了current-line-number或save-line-number参数，系统会跳过不相关的配置，从指定的行号开始查找两个配置文件的不同。系统比较出不同之处以后，将从两者有差异的地方开始显示字符，默认显示120个字符，如果从该不同之处到文件末尾不足120个字符，将显示到文件尾为止。

配置文件重置



功能	命令
清除下次启动时加载的配置文件	reset saved-configuration

```
<Huawei>reset saved-configuration
Warning: The action will delete the saved configuration in the
device.
The configuration will be erased to reconfigure. Continue? [Y/N]:y
Warning: Now clearing the configuration in the device.
Apr 10 2013 14:43:13-08:00 Huawei %%01CFM/4/RST_CFG(1)[2]:The user
chose Y when
deciding whether to reset the saved configuration.
Info: Succeeded in clearing the configuration in the device.
```




reset saved-configuration命令用来清除存储设备中启动配置文件的内容。

执行该命令后，如果不使用命令startup saved-configuration重新指定设备下次启动时使用的配置文件，也不使用save命令保存配置文件，则设备下次启动时会采用缺省的配置参数进行初始化。

存储设备

- SDRAM
- Flash
- NVRAM
- SD Card
- USB

```
<Huawei>display version
……
SDRAM Memory Size    : 2048     M bytes
Flash Memory Size    : 16       M bytes
NVRAM Memory Size    : 512      K bytes
SD Card1 Memory Size : 1882     M bytes
……
```

 Page 15



存储设备包括SDRAM、Flash、NVRAM、SD卡、U盘。例如，AR2200的路由器有内置的闪存和SD卡（槽位号为sd1）。该路由器提供了两个预留USB插槽(usb0 and usb1) 和一个SD卡插槽(sd0)。S5700交换机包含一个内置闪存，该闪存根据型号不同，存储容量也不同，S5700C-HI，S5700-LI，S5700S-LI和S5710-EI支持64M闪存，其他型号支持32M闪存。执行**display version**命令可以查看华为存储设备的详细信息。

存储设备修复

```
<Huawei>fixdisk flash:
Fixdisk flash: will take long time if needed
%Fixdisk flash: completed.

<Huawei>fixdisk sd1:
sd1:/  - disk check in progress
……
%Fixdisk sd1: completed.
```

- 当存储设备的文件系统出现异常时，可以通过fixdisk命令进行修复。



fixdisk命令用来对文件系统出现异常的存储设备进行修复。当存储设备上的文件系统出现异常时，终端会给出提示信息，此时建议使用此命令进行修复，但不确保修复成功。执行此命令后，如果仍然收到系统建议修复的信息，则表示物理介质可能已经损坏。

此命令是问题修复类命令，在系统未出现问题时，建议用户不要执行此命令。

存储设备格式化

```
<Huawei>format flash:
All data(include configuration and system startup file) on flash:
will be lost , proceed with format? (y/n)[n]:

<Huawei>format sdl:
All data(include configuration and system startup file) on sdl: will
be lost , proceed with format? (y/n)[n]:
```

- 格式化会导致数据丢失！



当文件系统出现异常无法修复时，并且确认不再需要存储器上的所有数据时，可格式化存储设备。格式化存储设备会导致设备上所有文件的丢失，且这些文件不能恢复。

format [devicename]命令用来格式化存储器。在执行**format**命令时，需要指定devicename参数，表示格式化特定的存储器。执行此命令后，会清空指定存储器中的所有文件和目录，并且不可恢复。请谨慎使用此命令！



总结

- 设备中的文件属性中有drw，其中d代表什么含义?
- 如果设备中有多个配置文件，如何指定下次启动时使用的配置文件?



1. d表明是个目录。r，w是可读出，可写入的意思。
2. 配置文件可以不使用默认的文件名vrp.cfg，而用其他指定的名称保存在路由器或者交换机中。如果需要指定某一配置文件为下次启动时使用的配置文件，可以执行**startup saved-configuration** [configuration-file-name]命令，这里的配置文件名包括文件名称和扩展名。


谢谢
www.huawei.com


VRP系统管理
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

为了满足企业业务对网络的需求，网络设备中的系统文件需要不断进行升级。另外，网络设备中的配置文件也需要时常进行备份，以防设备故障或其他灾害给业务带来损害。在升级和备份系统文件或配置文件时，经常会使用FTP和TFTP来传输文件。



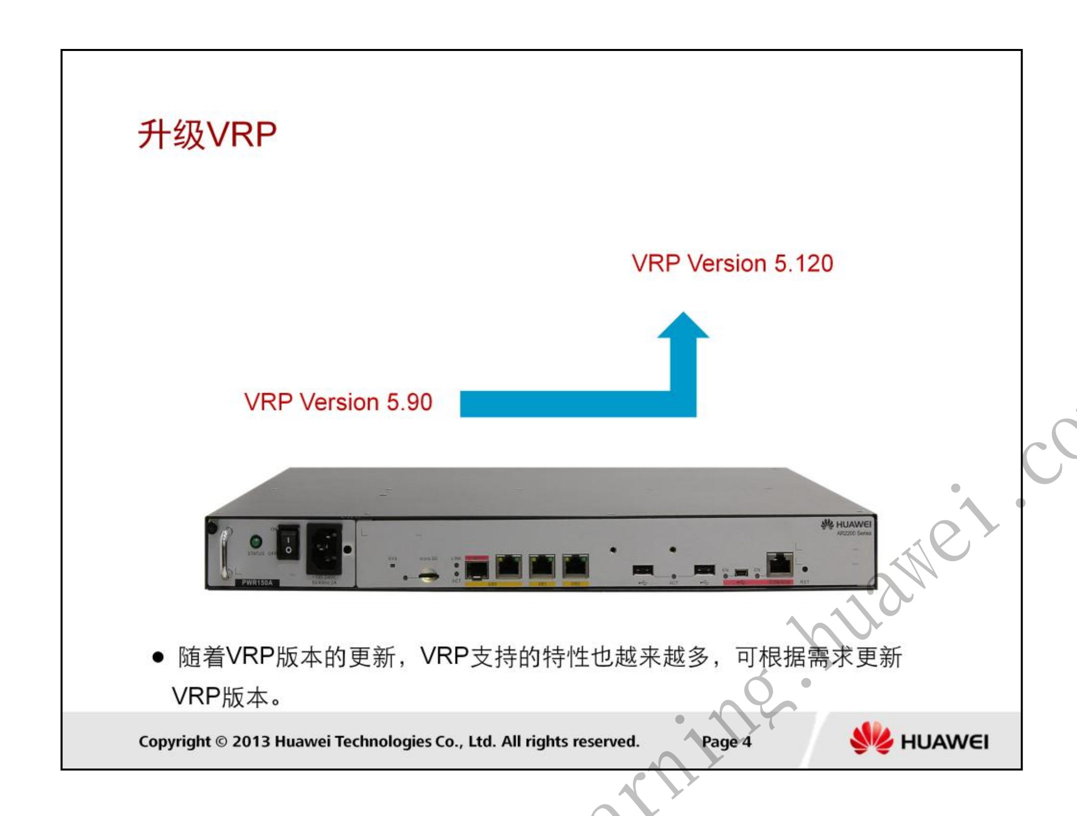


学习目标

学完本课程后，您应该能：

- 掌握FTP和TFTP的应用
- 掌握VRP升级的方法

HUAWEI

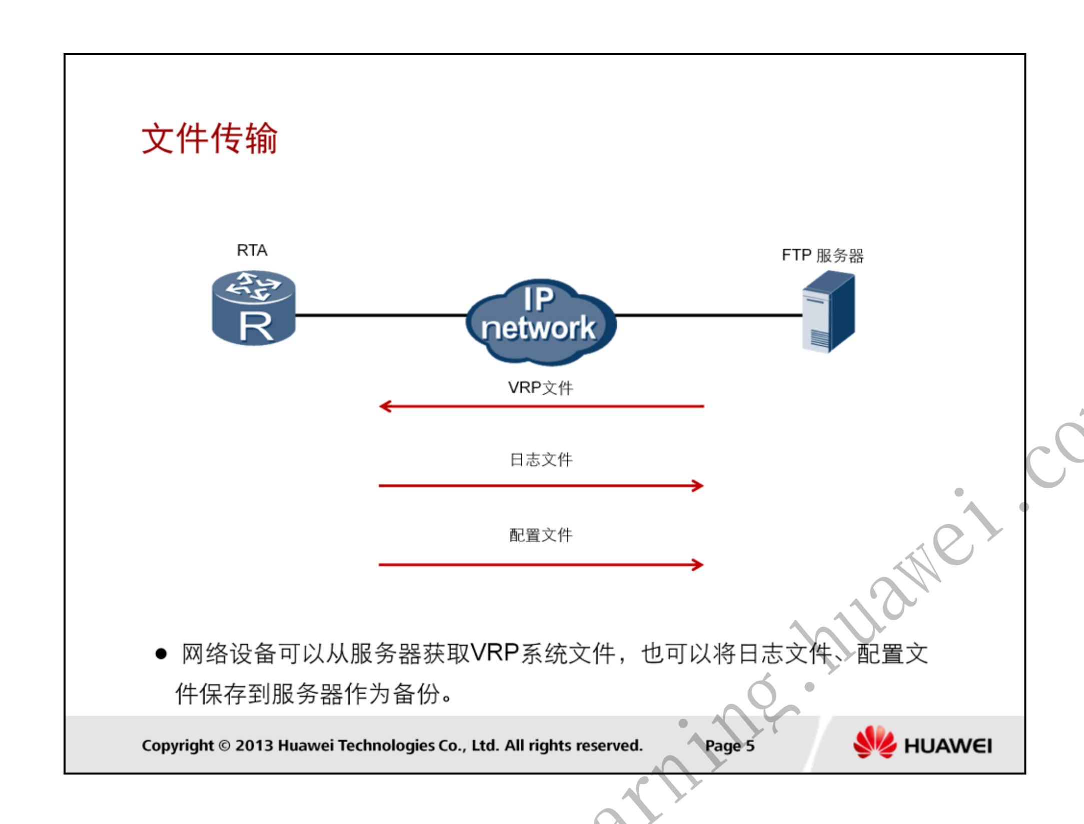


随着网络技术和应用的飞速发展，VRP也在不断的更新，VRP的命名由VRP自身版本号和关联产品版本号两部分组成。华为ARG3路由器和X7交换机使用的VRP版本为VRP5，VRP5可以和不同的产品版本相关联。随着产品版本增加，支持的特性也在增加。产品版本格式包含Vxxx（产品码），Rxxx(大版本号），Cxx(小版本号）。如果VRP产品版本有补丁，VRP产品版本号中还会包括SPC部分。

举例如下:

Version 5.90 (AR2200 V200R001C00)，VRP版本为5.90，产品版本号为V200R001C00。

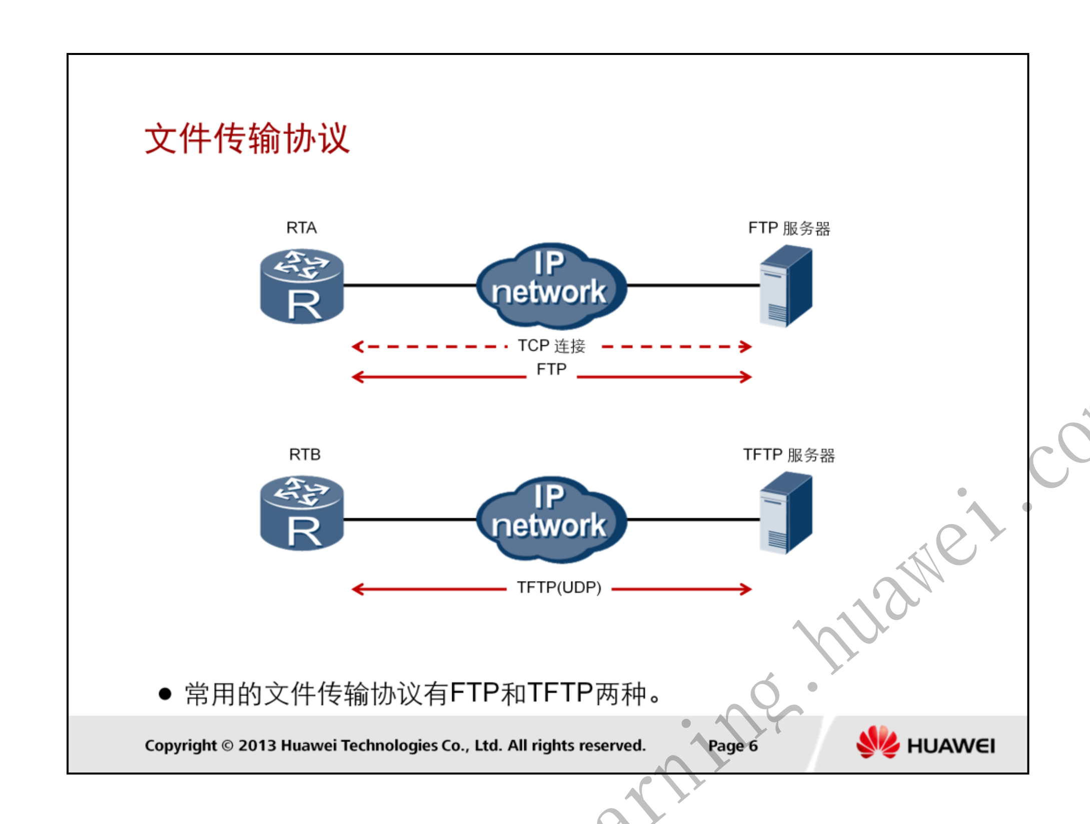
Version 5.120 (AR2200 V200R003C00SPC200)，VRP版本为5.120，产品版本号为V200R003C00SPC200，此产品版本包含有补丁包。



文件传输是指发送文件到远程服务器，或者从远程服务器获取文件的过程。

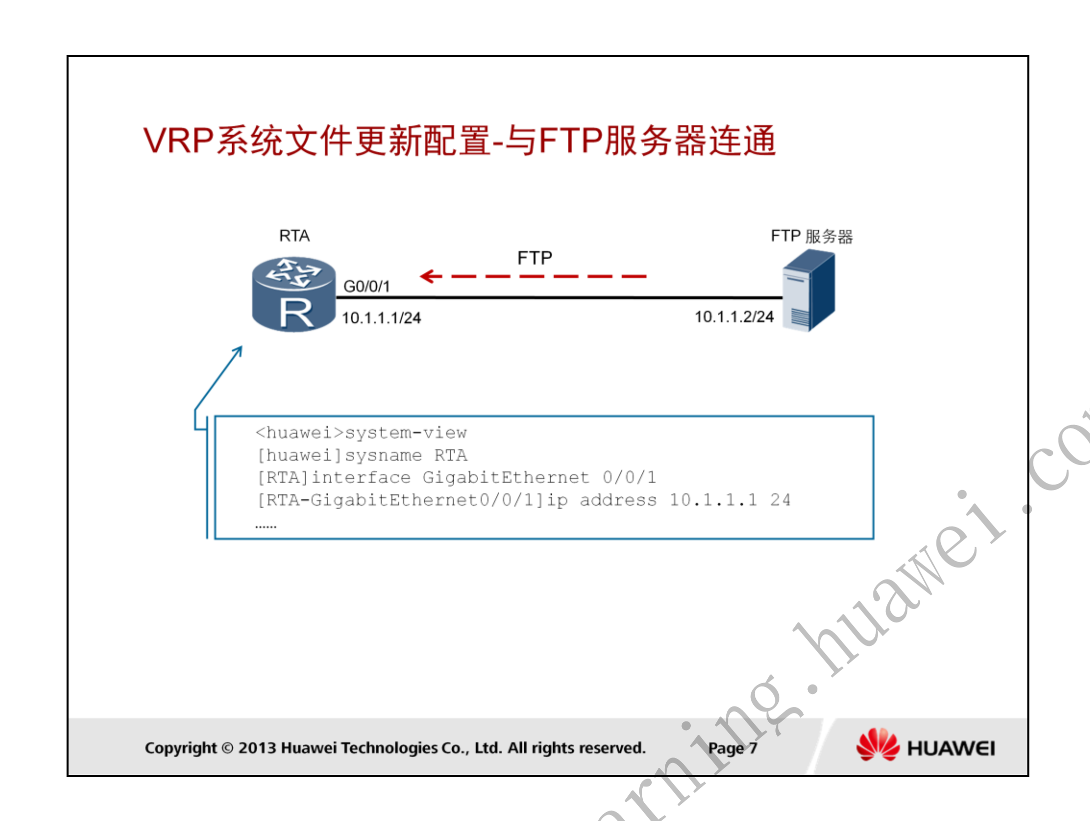
在实际场景中，为满足企业业务的需求，设备的VRP文件需要更新。如图中所示，设备在与服务器建立连接之后，可以从服务器获取新的VRP，完成更新工作。

为避免数据丢失对业务造成影响，设备的配置文件和日志文件也通常会进行远程备份。如图中所示，设备在与服务器建立连接之后，可以将配置文件和日志文件传输到服务器上，完成备份工作。当设备上的文件丢失后，可以恢复之前服务器上备份的配置文件和日志文件。

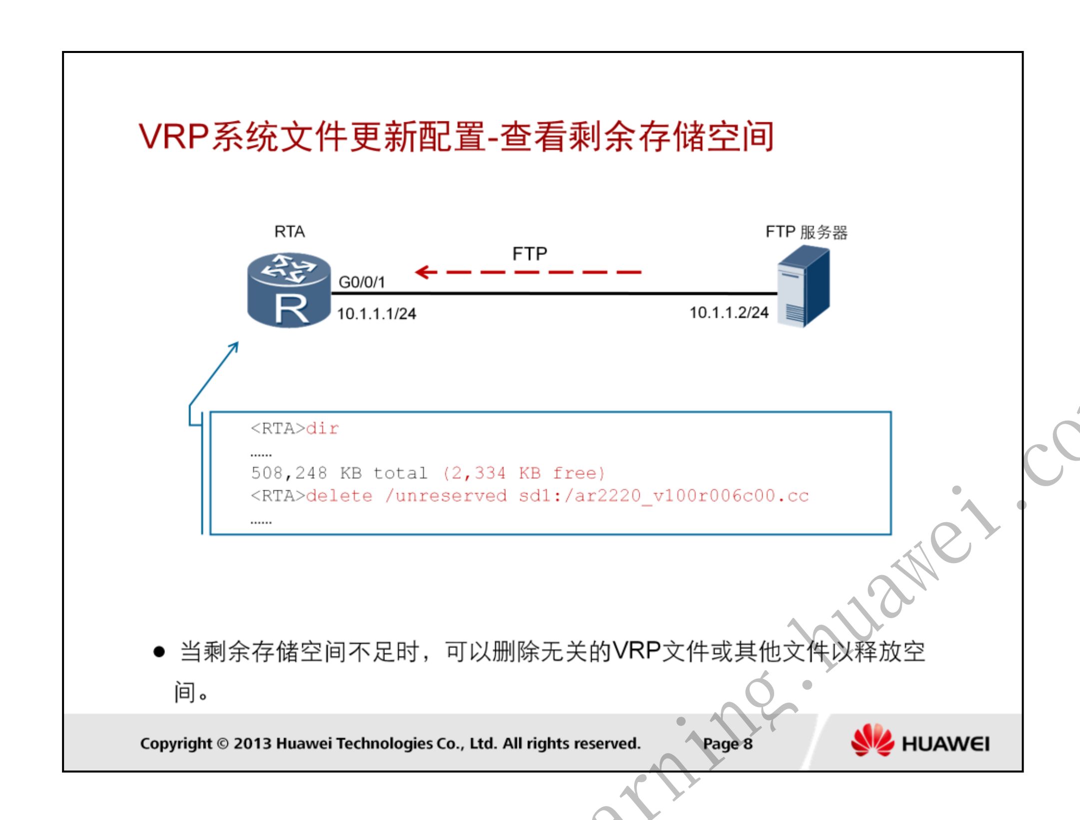


FTP（File Transfer Protocol）是TCP/IP协议族中的一种应用层协议，称为文件传输协议。FTP的主要功能是向用户提供本地和远程主机之间的文件传输。在进行版本升级、日志下载和配置保存等业务操作时，会广泛地使用到FTP。FTP采用两个TCP连接：控制连接和数据连接。其中控制连接用于连接控制端口，传输控制命令；数据连接用于连接数据端口，传输数据。在控制连接建立后，数据连接通过控制端口的命令建立起连接，进行数据的传输。FTP数据连接的建立有两种：主动模式和被动模式，两者的区别在于数据连接是由服务器发起还是由客户端发起。ARG3系列路由器既可以作为FTP Client又可以作为FTP Server。缺省情况下，AR2200采用主动模式建立数据连接。

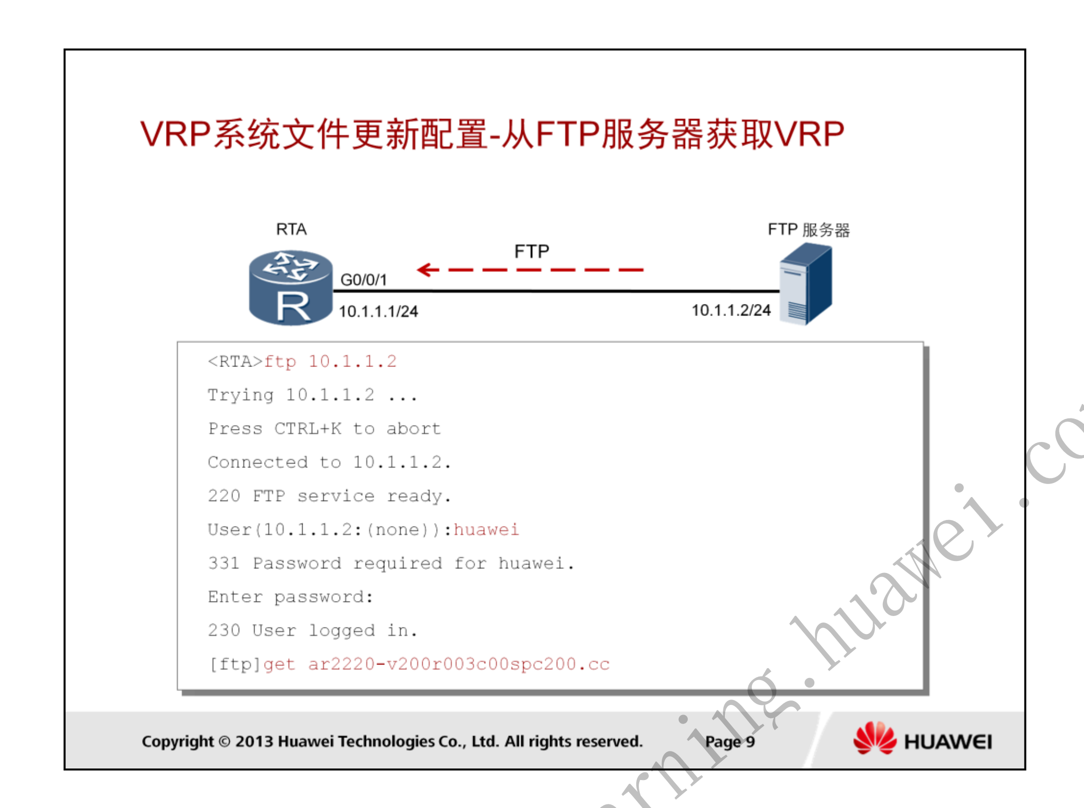
TFTP（Trivial File Transfer Protocol）是一种简化的文件传输协议。TFTP协议使用UDP协议进行文件的传输，由客户端发起TFTP传输请求，实现文件的上传和下载。ARG3系列路由器只可以作为TFTP客户端。



本例描述了华为ARG3路由器通过FTP获取VRP文件的过程。ARG3路由器作为FTP客户端，从FTP服务器中获取新的VRP文件并完成更新工作。在通过FTP传输任何数据前，首先必须保证FTP客户端和FTP服务器之间可以通信。



在获取VRP文件之前，还要确认设备有足够的存储空间来存储新获取的VRP文件。可以用dir命令来检查当前目录中的文件和可用空间。如果存储空间不足，需要删除无关文件来释放足够的空间。本例中通过使用delete命令删除已存在的无关VRP文件来获取了足够的空间。



如需从FTP服务器获取文件，客户端需使用**ftp** [*ip address*]命令来发起FTP连接请求，其中ip address指的是FTP服务器的IP地址。客户端和FTP服务器建立连接之后，客户端需要使用FTP服务器中配置的用户名和密码进行认证。认证通过后，客户端可以访问FTP服务器，并且能够查看和下载存储在服务器中的文件。

本例中客户端使用ftp 10.1.1.2和FTP服务器建立了连接，使用get ar2220-v200r003c00spc200.cc命令可以获取位于FTP服务器上的VRP文件。

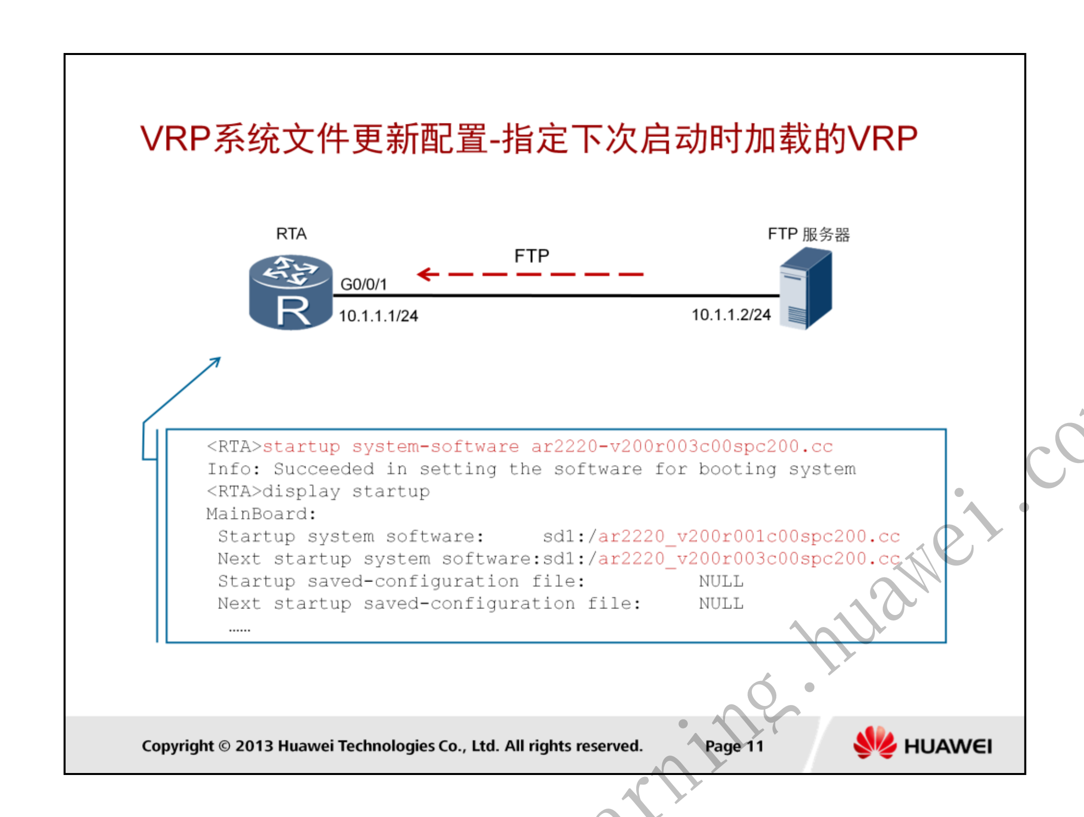
VRP系统文件更新配置-从TFTP服务器获取VRP


```
<RTA>tftp 10.1.1.2 get ar2220-v200r003c00spc200.cc
```

- ARG3路由器只需使用上述命令即可从TFTP服务器获取VRP文件。

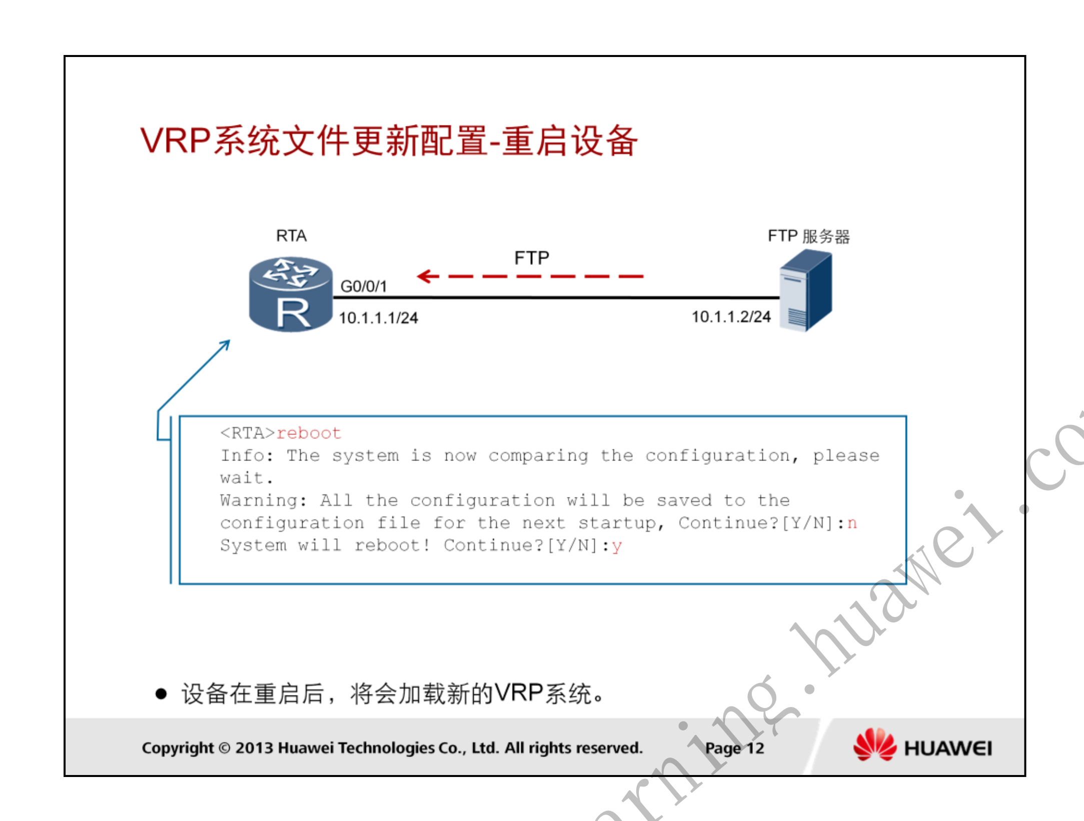
 Page 10 HUAWEI

如果客户端需要从TFTP服务器获取VRP文件，则不需要首先和TFTP服务器建立连接。ARG3系列路由器和X7系列交换机只能作为TFTP客户端。本例中客户端通过配置**tftp 10.1.1.2 get ar2220-v200r003c00spc200.cc**就可以从TFTP服务器获取VRP文件。



从服务器成功获取VRP文件后，还需要配置此文件为设备下次启动的系统文件，否则，设备仍会使用旧版本的VRP系统文件。在设备上通过使用**startup system-software**命令可以指定设备下次启动的系统文件。VRP系统文件必须存储在根目录，否则系统不能正常运行。

此例中可以使用**display startup**命令去验证系统启动文件是否已经变更，显示信息中Startup system software显示当前系统启动使用的VRP文件，Next startup system software显示下次系统启动使用的VRP文件。



确认系统下次启动软件正确后，需要重启设备。使用**reboot**命令可以重启设备。输入此命令后，系统会提示是否保存配置文件；实际中，可根据需要进行选择，本例中选择了不保存配置。



总结

- 设备作为FTP客户端时，如何从服务器下载VRP?
- 在完成VRP升级并重启之后，管理员如何确认升级成功?

 Page 13 

1. 首先必须保证客户端和服务器之间可以通信，然后客户端需使用**ftp**[*ip address*]命令与服务器建立FTP连接，建立连接之后需要输入正确的用户名和密码进行验证，验证通过后使用**get**命令即可下载VRP。
2. 管理员可以使用**display startup**命令验证系统启动的VRP软件，以此来判断VRP升级是否成功。


谢谢
www.huawei.com

Module-3

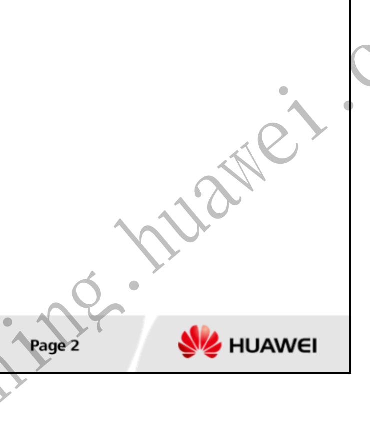
交换网的运行


交换网络基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

常见的以太网设备包括Hub、交换机等。交换机工作在数据链路层，它有效地隔离了以太网中的冲突域，极大地提升了以太网的性能。

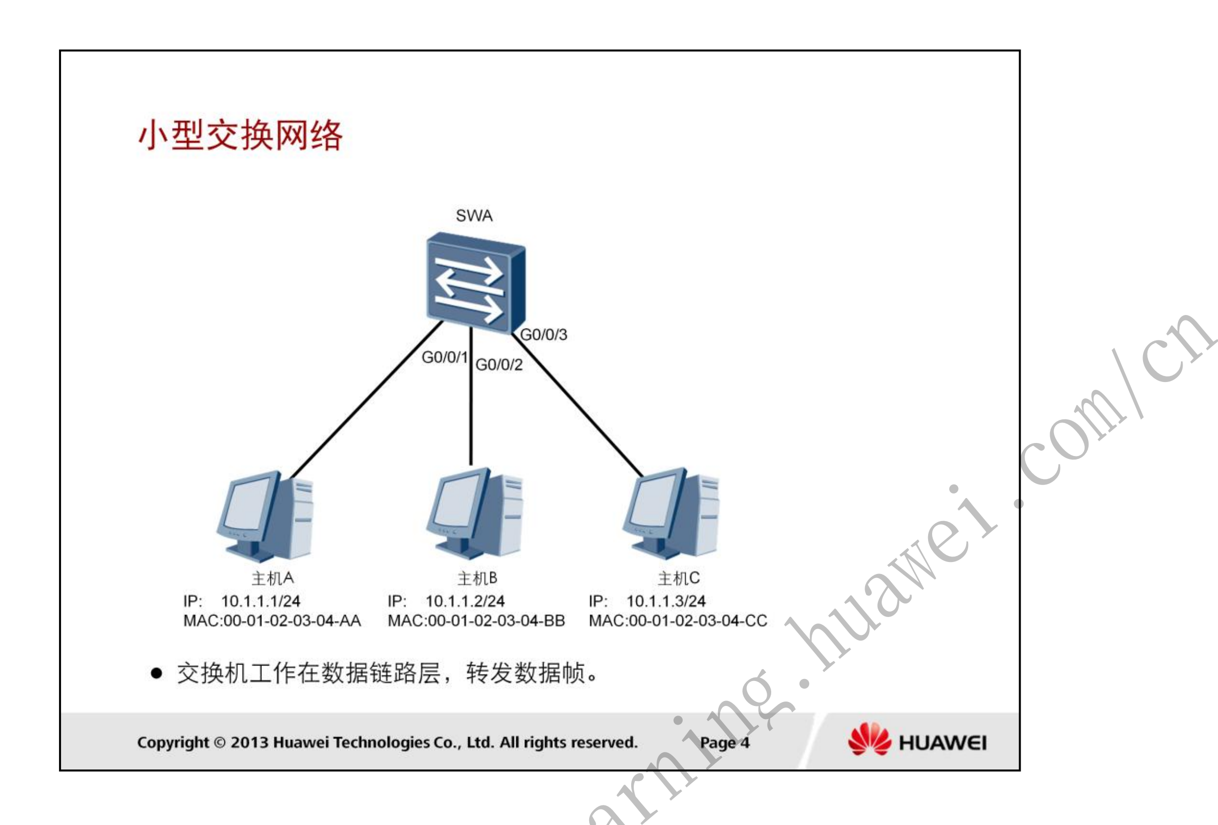




学习目标

学完本课程后，您应该能：

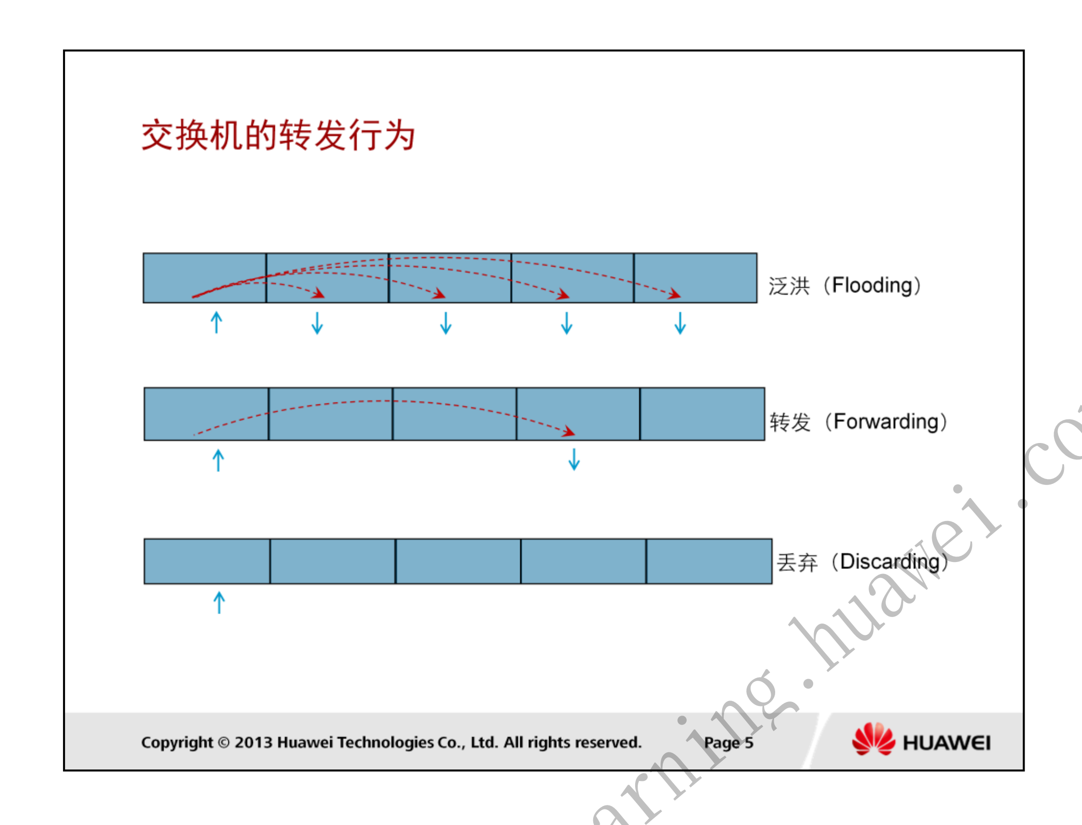
- 掌握交换机的基本工作原理
- 掌握交换机的基本配置



随着企业网络的发展，越来越多的用户需要接入到网络，交换机提供的大量的接入端口能够很好地满足这种需求。同时，交换机也彻底解决了困扰早期以太网的冲突问题，极大地提升了以太网的性能，同时也提高了以太网的安全性。

交换机工作在数据链路层，对数据帧进行操作。在收到数据帧后，交换机会根据数据帧的头部信息对数据帧进行转发。

接下来我们以小型交换网络为例，讲解交换机的基本工作原理。



交换机中有一个MAC地址表，里面存放了MAC地址与交换机端口的映射关系。MAC地址表也称为CAM（Content Addressable Memory）表。

如图所示，交换机对帧的转发操作行为一共有三种：放洪（Flooding），转发（Forwarding），丢弃（Discarding）。

1. 泛洪：交换机把从某一端口进来的帧通过所有其它的端口转发出去（注意，“所有其它的端口”是指除了这个帧进入交换机的那个端口以外的所有端口）。
2. 转发：交换机把从某一端口进来的帧通过另一个端口转发出去（注意，“另一个端口”不能是这个帧进入交换机的那个端口）。
3. 丢弃：交换机把从某一端口进来的帧直接丢弃。

交换机的基本工作原理可以概括地描述如下：

1. 如果进入交换机的是一个单播帧，则交换机会去MAC地址表中查找这个帧的目的MAC地址。

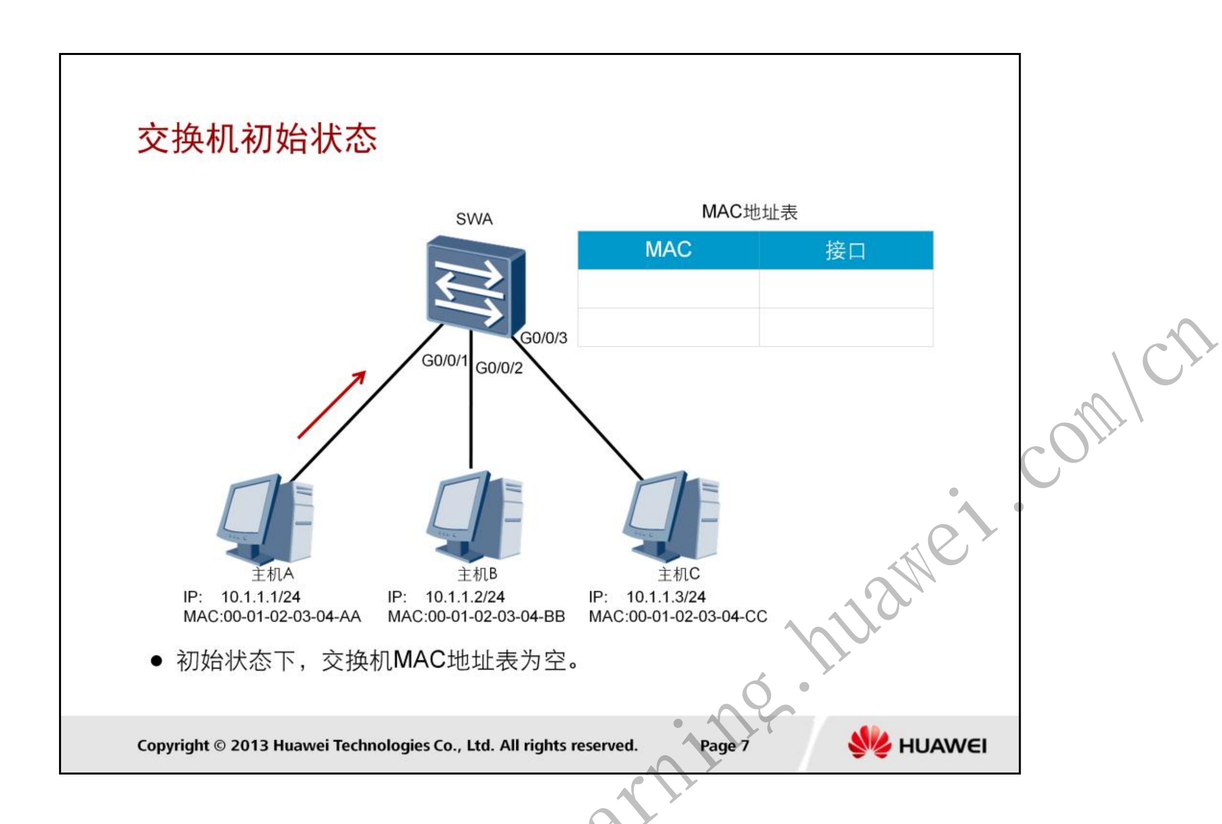
   1）如果查不到这个MAC地址，则交换机执行泛洪操作。

   2）如果查到了这个MAC地址，则比较这个MAC地址在MAC地址表中对应的端口是不是这个帧进入交换机的那个端口。如果不是，则交换机执行转发操作。如果是，则交换机执行丢弃操作。

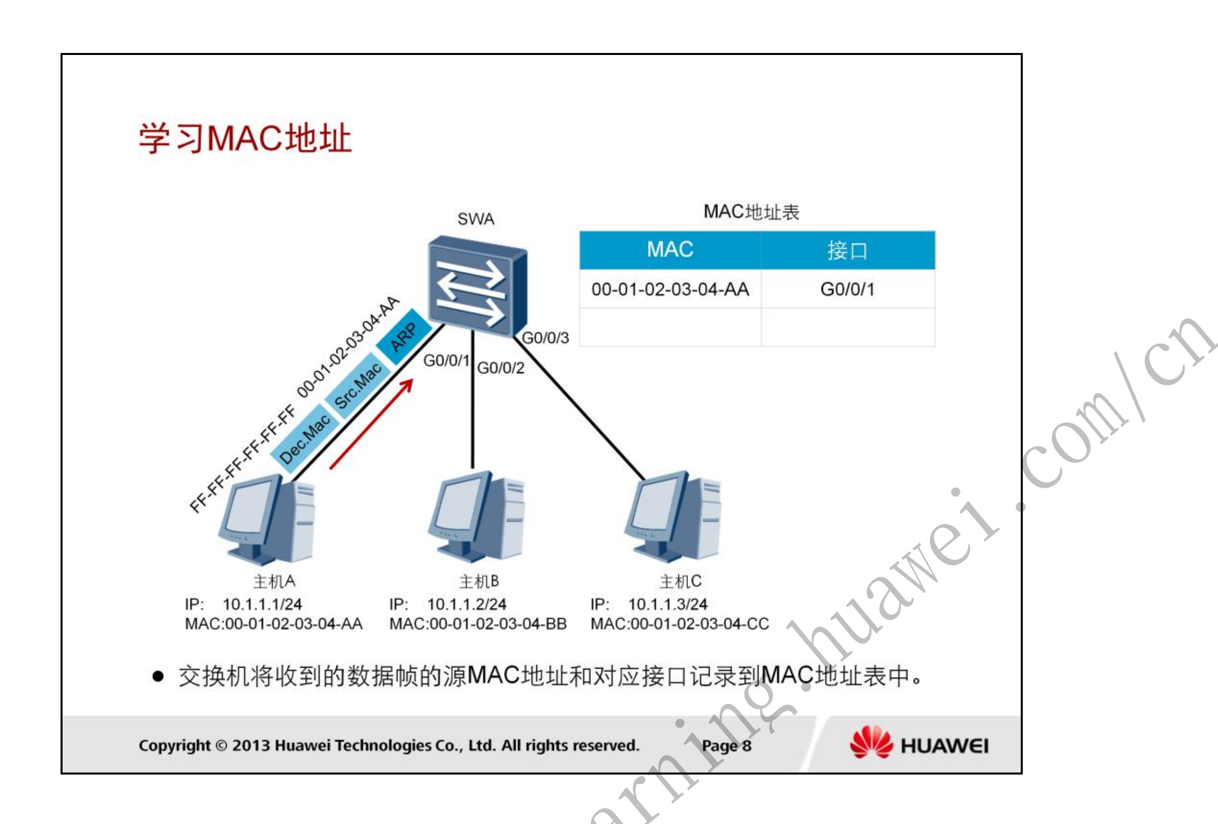
2. 如果进入交换机的是一个广播帧，则交换机不会去查MAC地址表，而是直接执行泛洪操作。

3. 如果进入交换机的是一个组播帧，则交换机的处理行为比较复杂，超出了这里的学习范围，所以略去不讲。

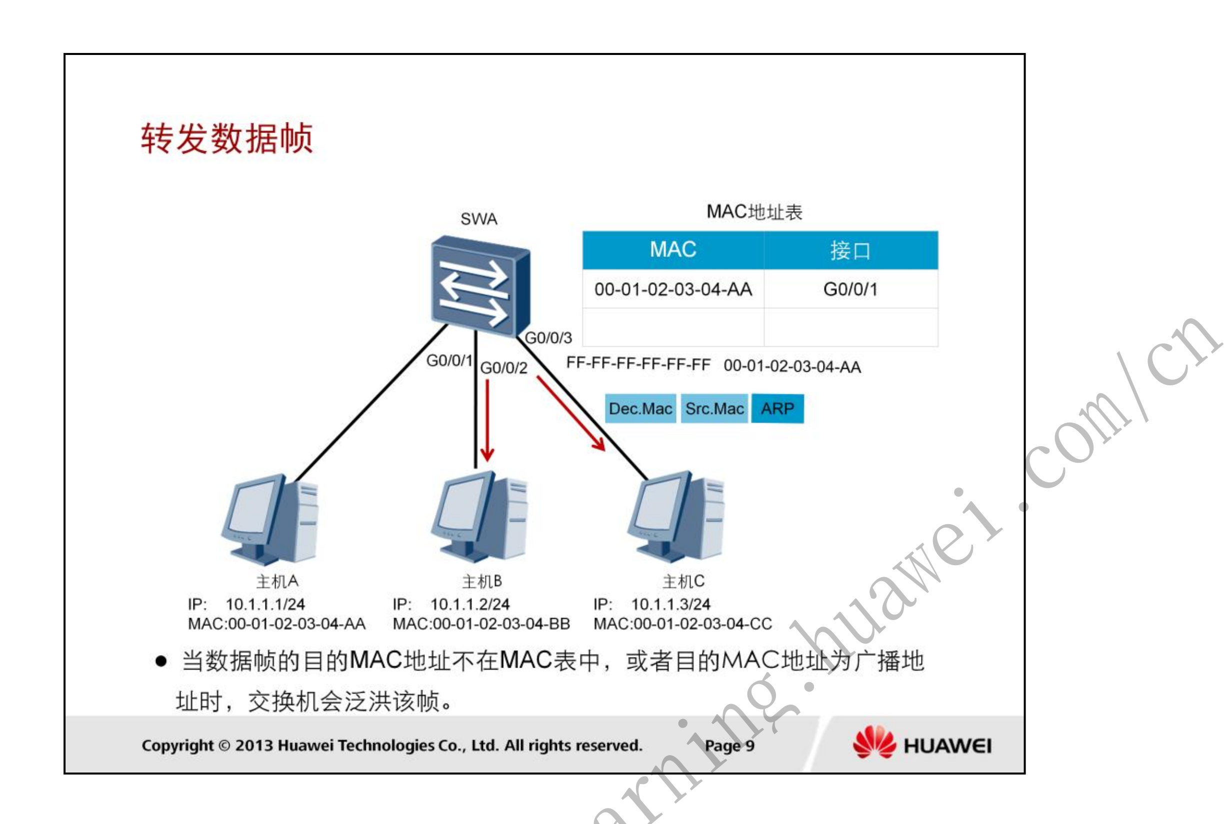
另外，交换机还具有学习能力。当一个帧进入交换机后，交换机会检查这个帧的源MAC地址，并将该源MAC地址与这个帧进入交换机的那个端口进行映射，然后将这个映射关系存放进MAC地址表。



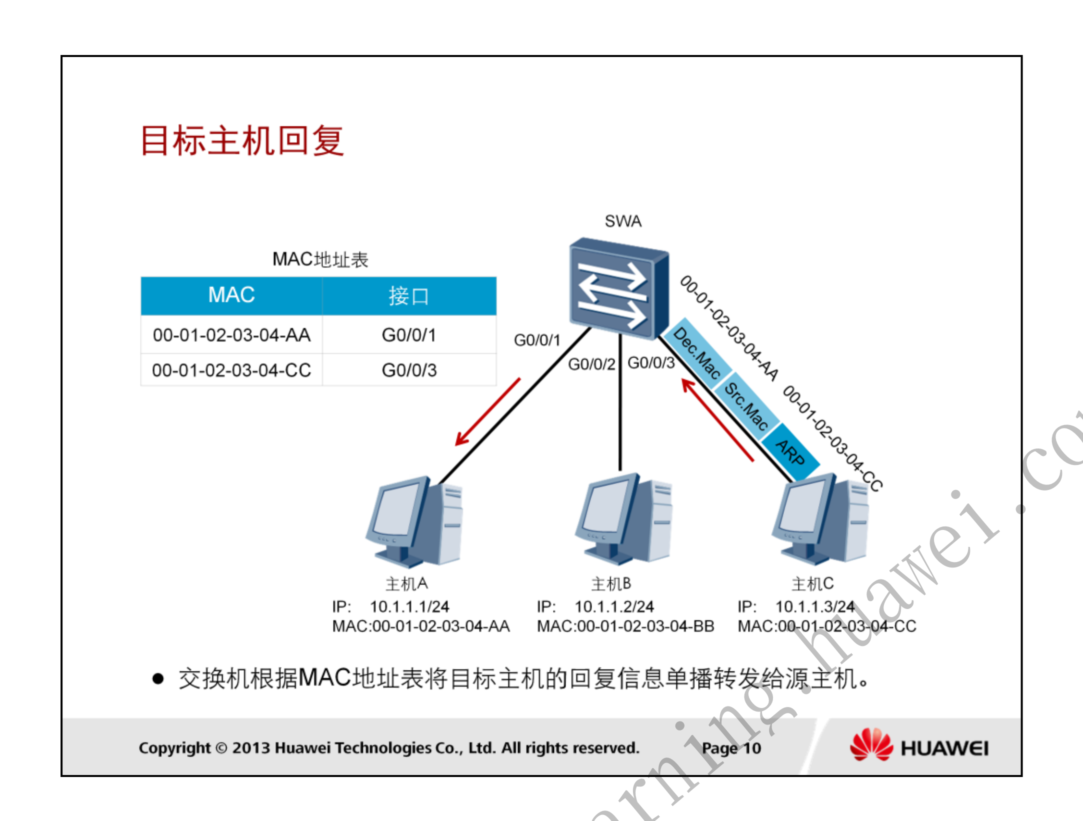
初始状态下，交换机并不知道所连接主机的MAC地址，所以MAC地址表为空。本例中，SWA为初始状态，在收到主机A发送的数据帧之前，MAC地址表中没有任何表项。



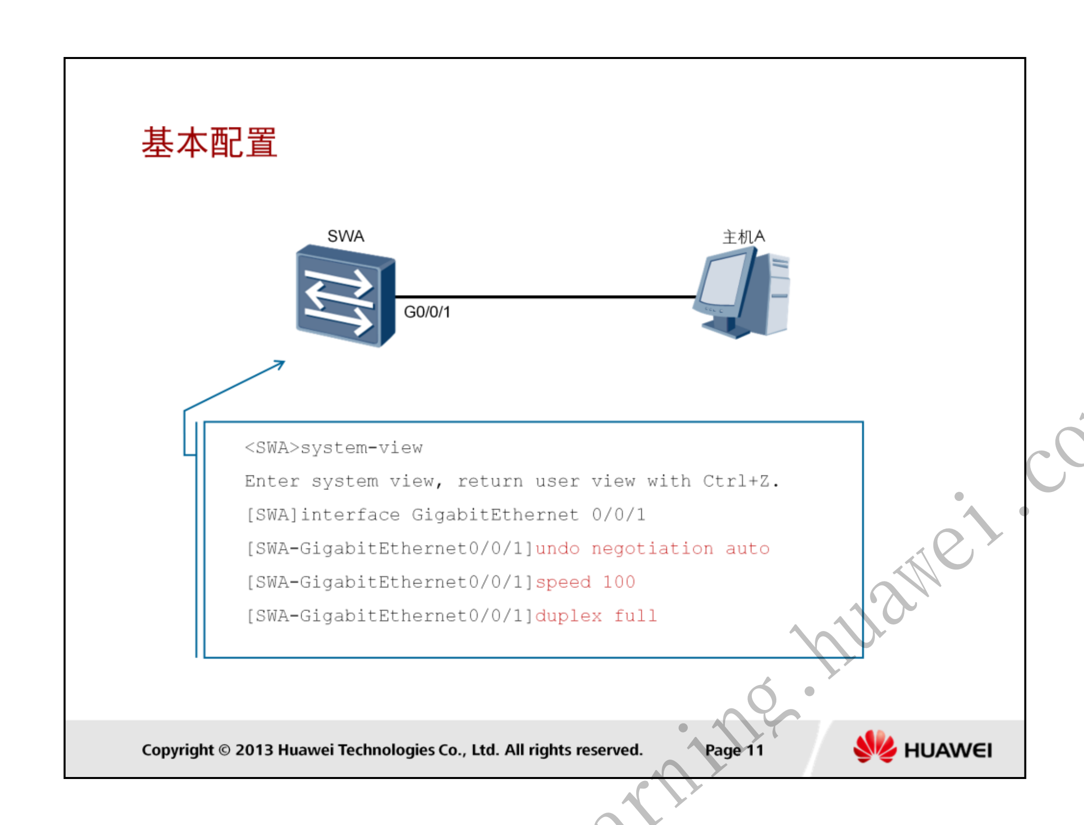
主机A发送数据给主机C时，一般会首先发送ARP请求来获取主机C的MAC地址，此ARP请求帧中的目的MAC地址是广播地址，源MAC地址是自己的MAC地址。SWA收到该帧后，会将源MAC地址和接收端口的映射关系添加到MAC地址表中。缺省情况下，X7系列交换机学习到的MAC地址表项的老化时间为300秒。如果在老化时间内再次收到主机A发送的数据帧，SWA中保存的主机A的MAC地址和G0/0/1的映射的老化时间会被刷新。此后，如果交换机收到目标MAC地址为00-01-02-03-04-AA的数据帧时，都将通过G0/0/1端口转发。



本例中主机A发送的数据帧的目的MAC地址为广播地址，所以交换机会将此数据帧通过G0/0/2和G0/0/3端口广播到主机B和主机C。



主机B和主机C接收到此数据帧后，都会查看该ARP数据帧。但是主机B不会回复该帧，主机C会处理该帧并发送ARP回应，此回复数据帧的目的MAC地址为主机A的MAC地址，源MAC地址为主机C的MAC地址。SWA收到回复数据帧时，会将该帧的源MAC地址和接口的映射关系添加到MAC地址表中。如果此映射关系在MAC地址表已经存在，则会被刷新。然后SWA查询MAC地址表，根据帧的目的MAC地址找到对应的转发端口后，从G0/0/1转发此数据帧。

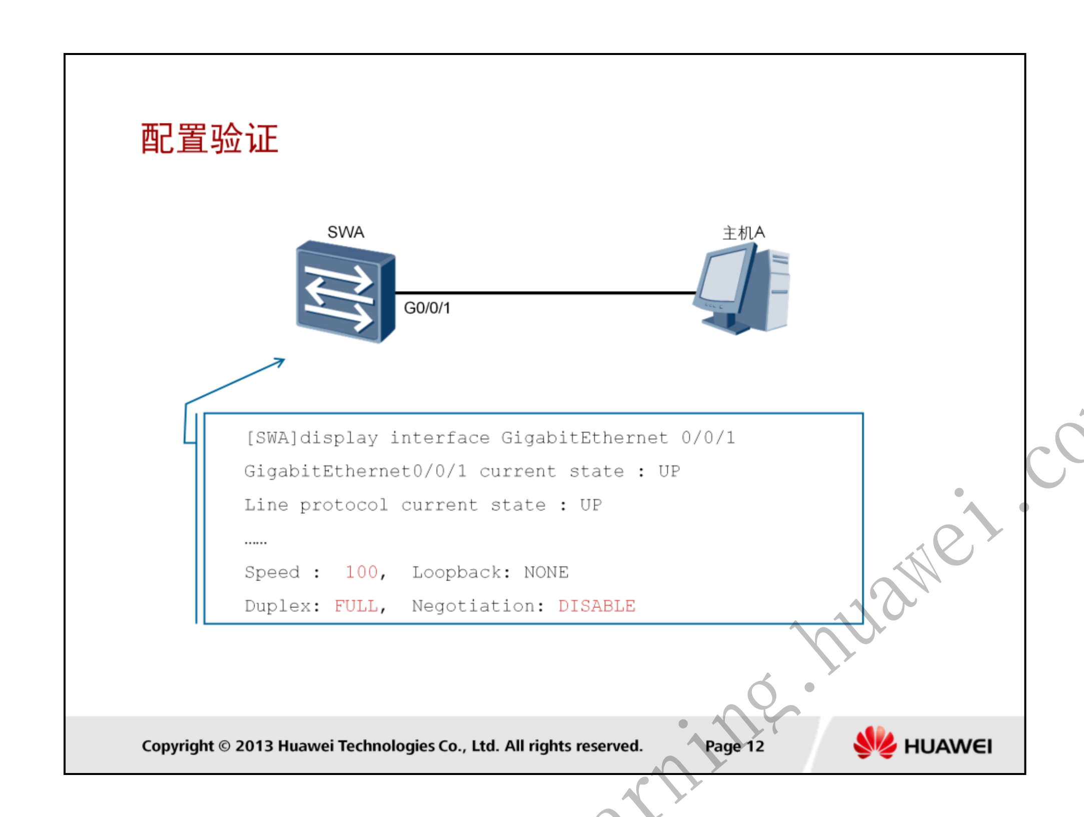


早期的以太网的工作模式都是10M半双工的。随着技术的发展，出现了全双工模式，接着又出现了百兆和千兆以太网。采用不同工作模式的设备无法直接相互通信；自协商技术的出现解决了不同以太网工作模式之间的兼容性问题。自协商的内容主要包括双工模式和运行速率。一旦协商通过，链路两端的设备就具有相同的工作参数。

negotiation auto命令用来设置以太网端口的自协商功能。端口是否应该使能自协商模式，要考虑对接双方设备的端口是否都支持自动协商。如果对端设备的以太网端口不支持自协商模式，则需要在本端端口上先使用**undo negotiation auto**命令配置为非自协商模式。之后，修改本端端口的速率和双工模式保持与对端一致，确保通信正常。

duplex命令用来设置以太网端口的双工模式。GE电口工作速率为1000Mbit/s时，只支持全双工模式，不需要与链路对端的端口共同协商双工模式。

speed命令用来设置端口的工作速率。配置端口的速率和双工模式之前需要先配置端口为非自协商模式。



display interface [*interface-type* [*interface-number* [.*subnumber*]]] 命令用来查看端口当前运行状态和统计信息。

current state表示端口的物理状态，如果为UP，表示端口处于打开状态。

Line protocol current state表示端口的链路协议状态，如果为UP，表示端口的链路协议处于正常的启动状态。

Speed表示端口的工作速率，SWA的G0/0/1端口工作速率为100Mbit/s。

Duplex表示端口的双工模式，SWA的G0/0/1端口双工模式为全双工。



总结

- 当一台主机从交换机的一个端口移动到另外一个端口时，交换机的MAC地址表会发生什么变化?





1. 当一台主机从交换机的一个端口移除时，交换机检测到物理链路Down，因此会从MAC地址表中清除对应主机的MAC表项。一旦主机连接到交换机另外一个端口，交换机会检测到新端口对应的物理链路UP。主机发送报文后，交换机就会学习到主机的MAC地址和新端口的映射关系，并且添加到MAC地址表中。


谢谢
www.huawei.com


STP原理与配置
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

为了提高网络可靠性，交换网络中通常会使用冗余链路。然而，冗余链路会给交换网络带来环路风险，并导致广播风暴以及MAC地址表不稳定等问题，进而会影响到用户的通信质量。生成树协议STP（Spanning Tree Protocol）可以在提高可靠性的同时又能避免环路带来的各种问题。





学习目标

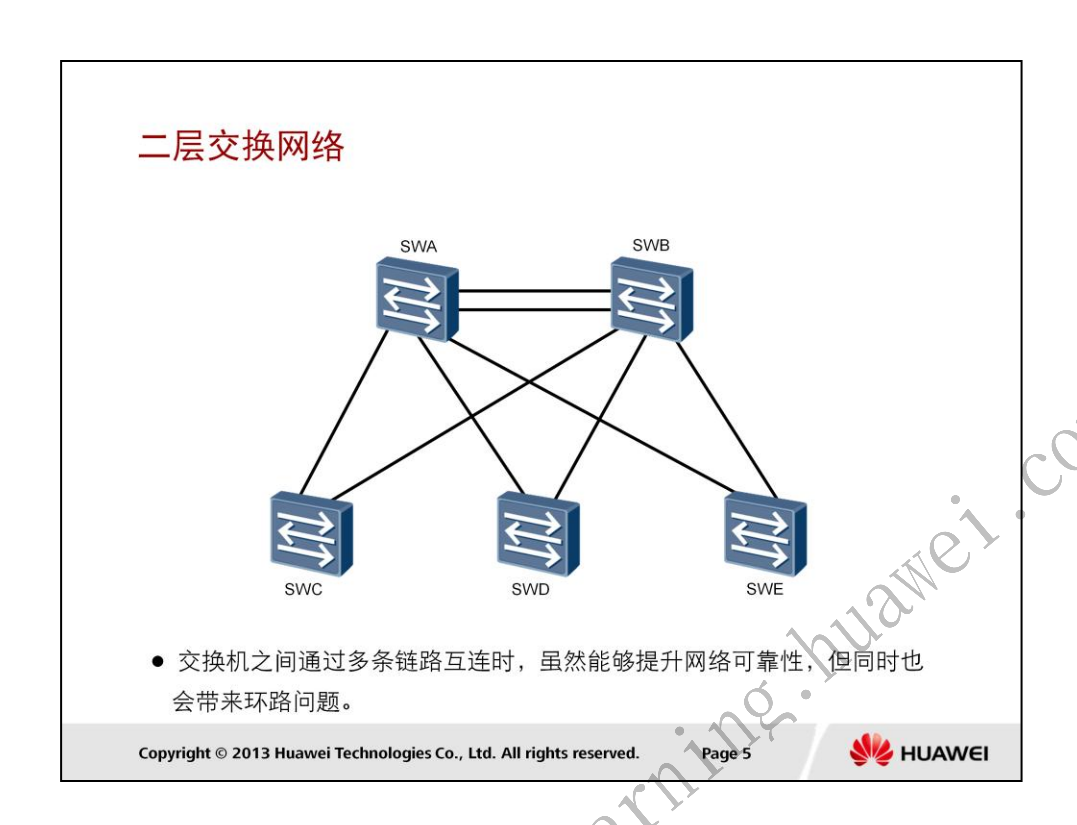
学完本课程后，您应该能：

- 掌握STP的工作原理
- 掌握STP的基本配置



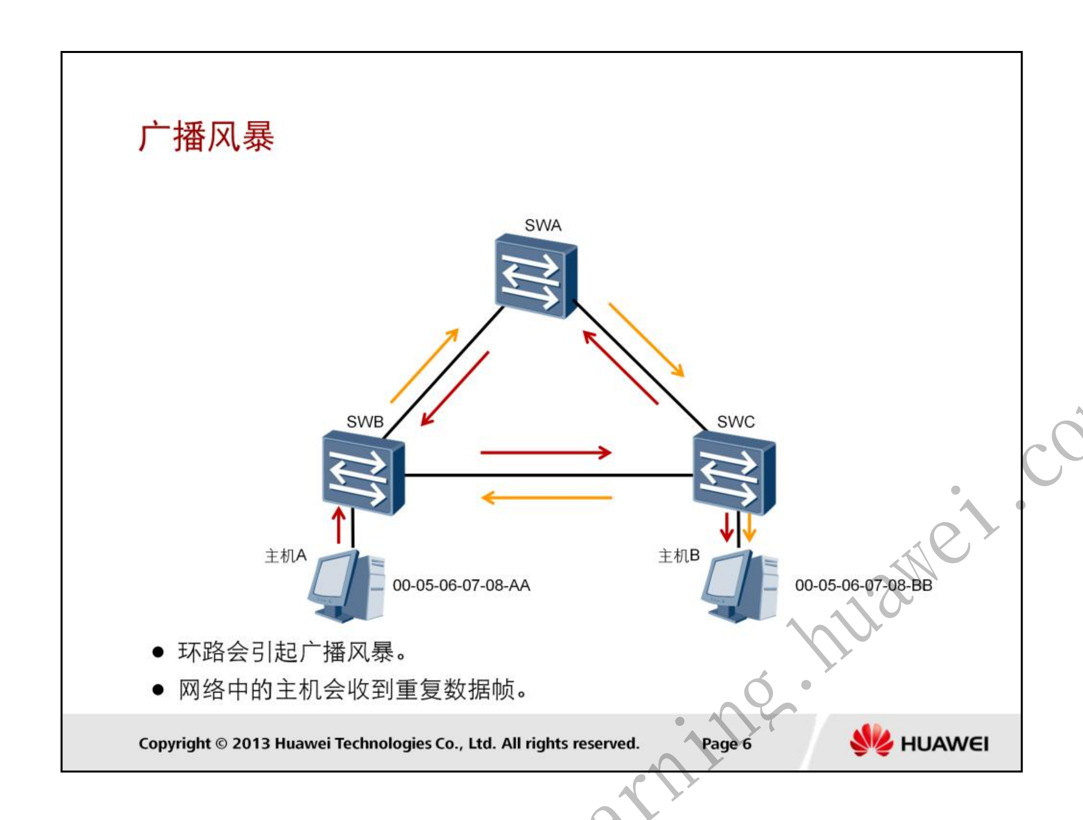
1. **环路引起的问题**
2. STP工作原理
3. STP拓扑变化
4. STP的配置

HUAWEI



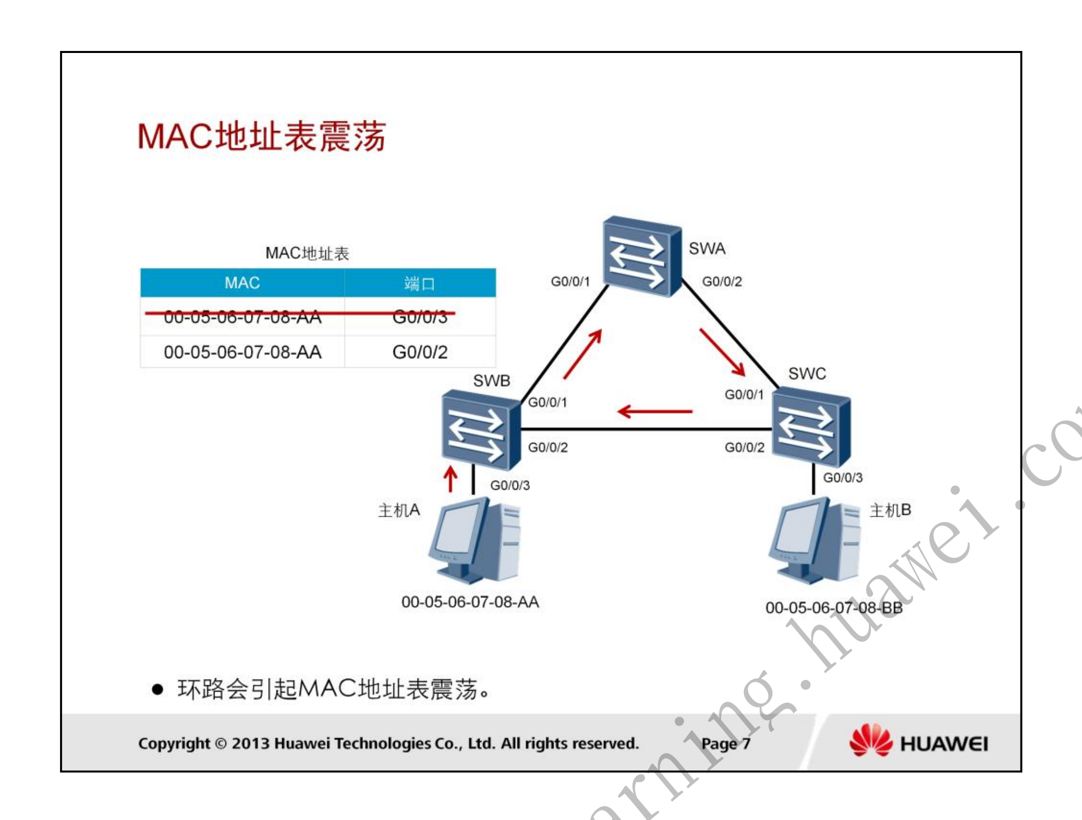
随着局域网规模的不断扩大，越来越多的交换机被用来实现主机之间的互连。如果交换机之间仅使用一条链路互连，则可能会出现单点故障，导致业务中断。为了解决此类问题，交换机在互连时一般都会使用冗余链路来实现备份。

冗余链路虽然增强了网络的可靠性，但是也会产生环路，而环路会带来一系列的问题，继而导致通信质量下降和通信业务中断等问题。



根据交换机的转发原则，如果交换机从一个端口上接收到的是一个广播帧，或者是一个目的MAC地址未知的单播帧，则会将这个帧向除源端口之外的所有其他端口转发。如果交换网络中有环路，则这个帧会被无限转发，此时便会形成广播风暴，网络中也会充斥着重复的数据帧。

本例中，主机A向外发送了一个单播帧，假设此单播帧的目的MAC地址在网络中所有交换机的MAC地址表中都暂时不存在。SWB接收到此帧后，将其转发到SWA和SWC，SWA和SWC也会将此帧转发到除了接收此帧的其他所有端口，结果此帧又会被再次转发给SWB，这种循环会一直持续，于是便产生了广播风暴。交换机性能会因此急速下降，并会导致业务中断。



交换机是根据所接收到的数据帧的源地址和接收端口生成MAC地址表项的。

主机A向外发送一个单播帧，假设此单播帧的目的MAC地址在网络中所有交换机的MAC地址表中都暂时不存在。SWB收到此数据帧之后，在MAC地址表中生成一个MAC地址表项，00-01-02-03-04-AA，对应端口为G0/0/3，并将其从G0/0/1和G0/0/2端口转发。此例仅以SWB从G0/0/1端口转发此帧为例进行说明。

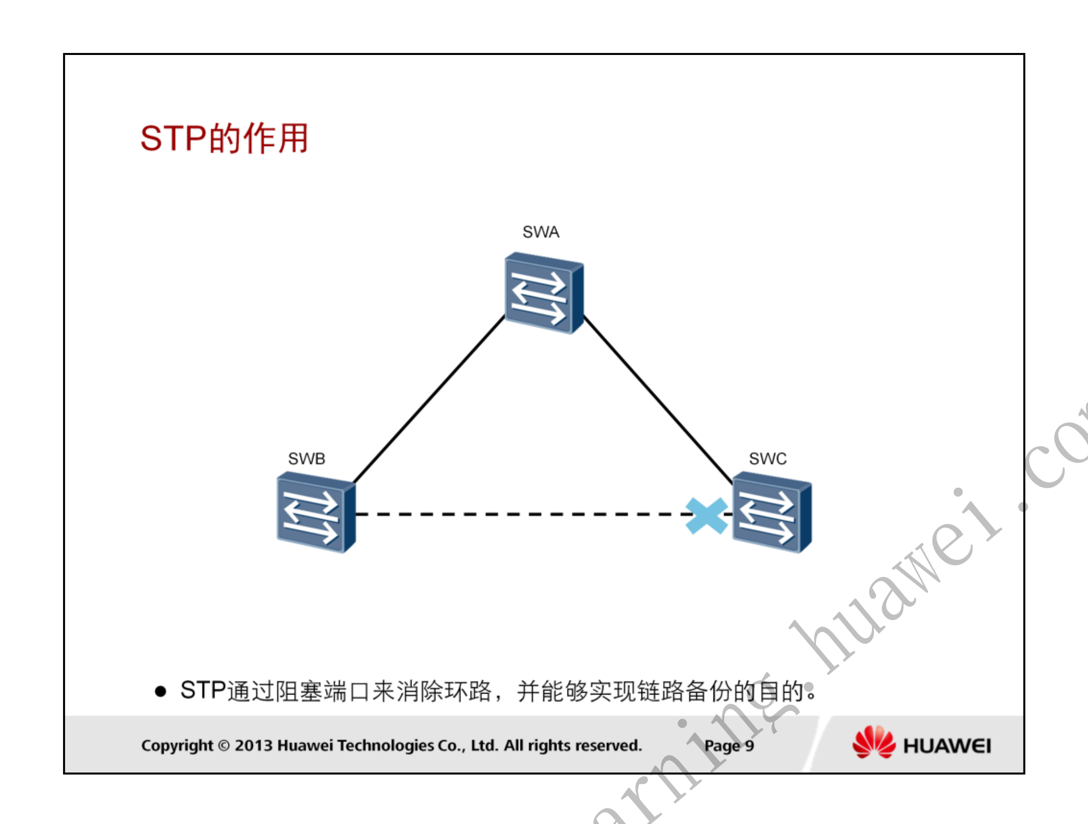
SWA接收到此帧后，由于MAC地址表中没有对应此帧目的MAC地址的表项，所以SWA会将此帧从G0/0/2转发出去。

SWC接收到此帧后，由于MAC地址表中也没有对应此帧目的MAC地址的表项，所以SWC会将此帧从G0/0/2端口发送回SWB，也会发给主机B。

SWB从G0/0/2接口接收到此数据帧之后，会在MAC地址表中删除原有的相关表项，生成一个新的表项，00-01-02-03-04-AA，对应端口为G0/0/2。此过程会不断重复，从而导致MAC地址表震荡。



目录



在以太网中，二层网络的环路会带来广播风暴，MAC地址表震荡，重复数据帧等问题，为解决交换网络中的环路问题，提出了STP。

STP的主要作用：

1. 消除环路：通过阻断冗余链路来消除网络中可能存在的环路。
2. 链路备份：当活动路径发生故障时，激活备份链路，及时恢复网络连通性。

STP操作

SWA
Root
D
D
SWB
R
D
R SWC
A
Non-Root
Non-Root

1. 选举一个根桥。
2. 每个非根交换机选举一个根端口。
3. 每个网段选举一个指定端口。
4. 阻塞非根、非指定端口。

 Page 10 HUAWEI

STP通过构造一棵树来消除交换网络中的环路。

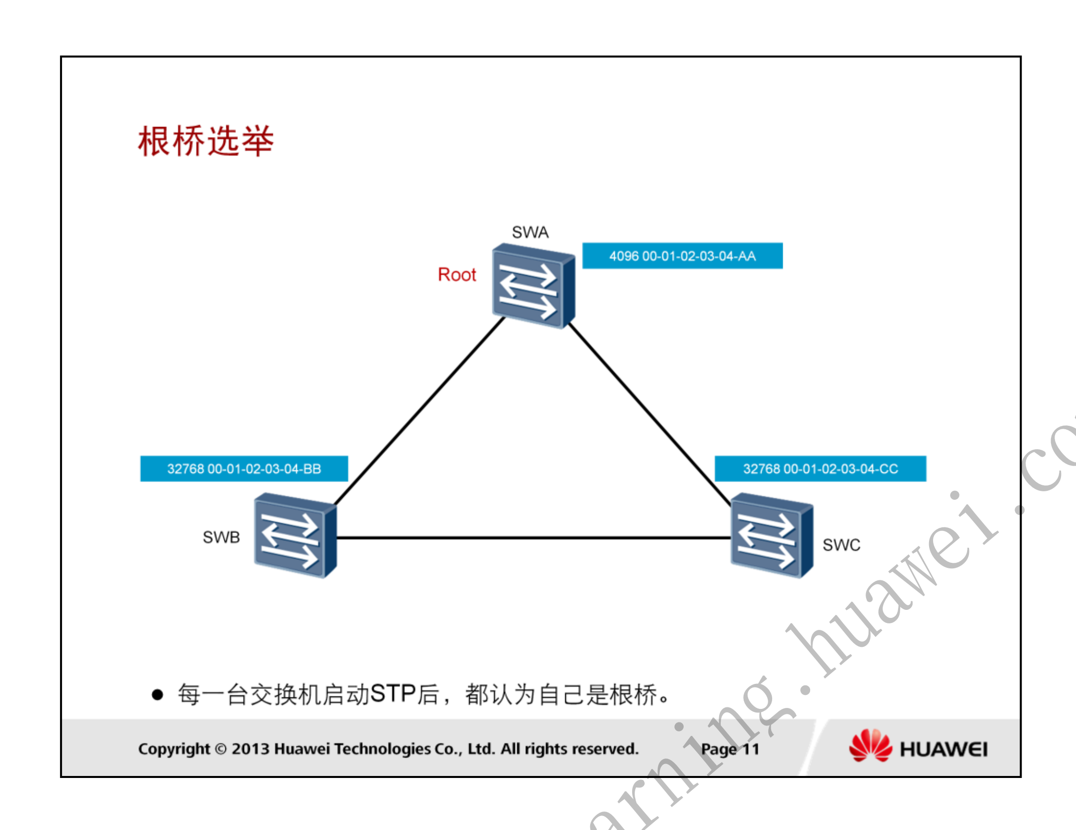
每个STP网络中，都会存在一个根桥，其他交换机为非根桥。根桥或者根交换机位于整个逻辑树的根部，是STP网络的逻辑中心，非根桥是根桥的下游设备。当现有根桥产生故障时，非根桥之间会交互信息并重新选举根桥，交互的这种信息被称为BPDU。BPDU中包含交换机在参加生成树计算时的各种参数信息，后面会有详细介绍。

STP中定义了三种端口角色：指定端口，根端口和预备端口。

指定端口是交换机向所连网段转发配置BPDU的端口，每个网段有且只能有一个指定端口。一般情况下，根桥的每个端口总是指定端口。

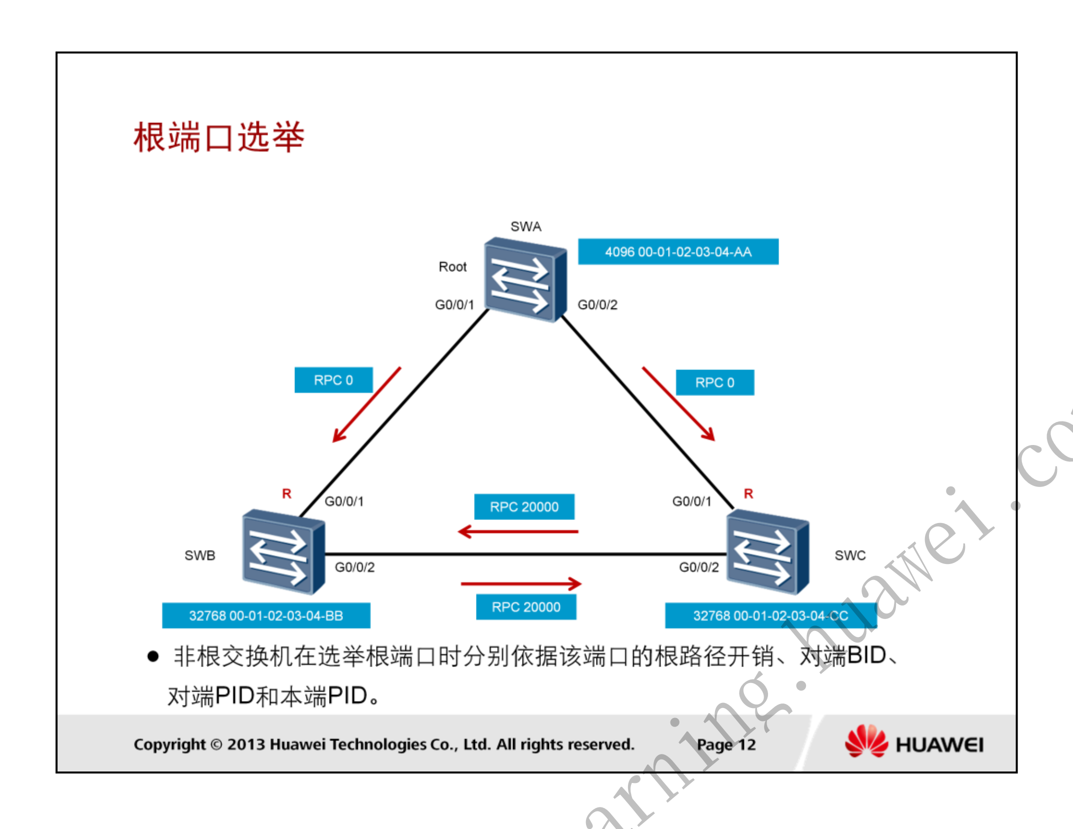
根端口是非根交换机去往根桥路径最优的端口。在一个运行STP协议的交换机上最多只有一个根端口，但根桥上没有根端口。

如果一个端口既不是指定端口也不是根端口，则此端口为预备端口。预备端口将被阻塞。



STP中根桥的选举依据的是桥ID，STP中的每个交换机都会有一个桥ID(Bridge ID) 。桥ID由16位的桥优先级（Bridge Priority）和48位的MAC地址构成。在STP网络中，桥优先级是可以配置的，取值范围是0～65535，默认值为32768。优先级最高的设备（桥ID最小）会被选举为根桥。如果优先级相同，则会比较MAC地址，MAC地址越小则越优先。

交换机启动后就自动开始进行生成树收敛计算。默认情况下，所有交换机启动时都认为自己是根桥，自己的所有端口都为指定端口，这样BPDU报文就可以通过所有端口转发。对端交换机收到BPDU报文后，会比较BPDU中的根桥ID和自己的桥ID。如果收到的BPDU报文中的桥ID优先级低，接收交换机会继续通告自己的配置BPDU报文给邻居交换机。如果收到的BPDU报文中的桥ID优先级高，则交换机会修改自己的BPDU报文的根桥ID字段，宣告新的根桥。



非根交换机在选举根端口时分别依据该端口的根路径开销、对端BID（Bridge ID）、对端PID（Port ID）和本端PID。

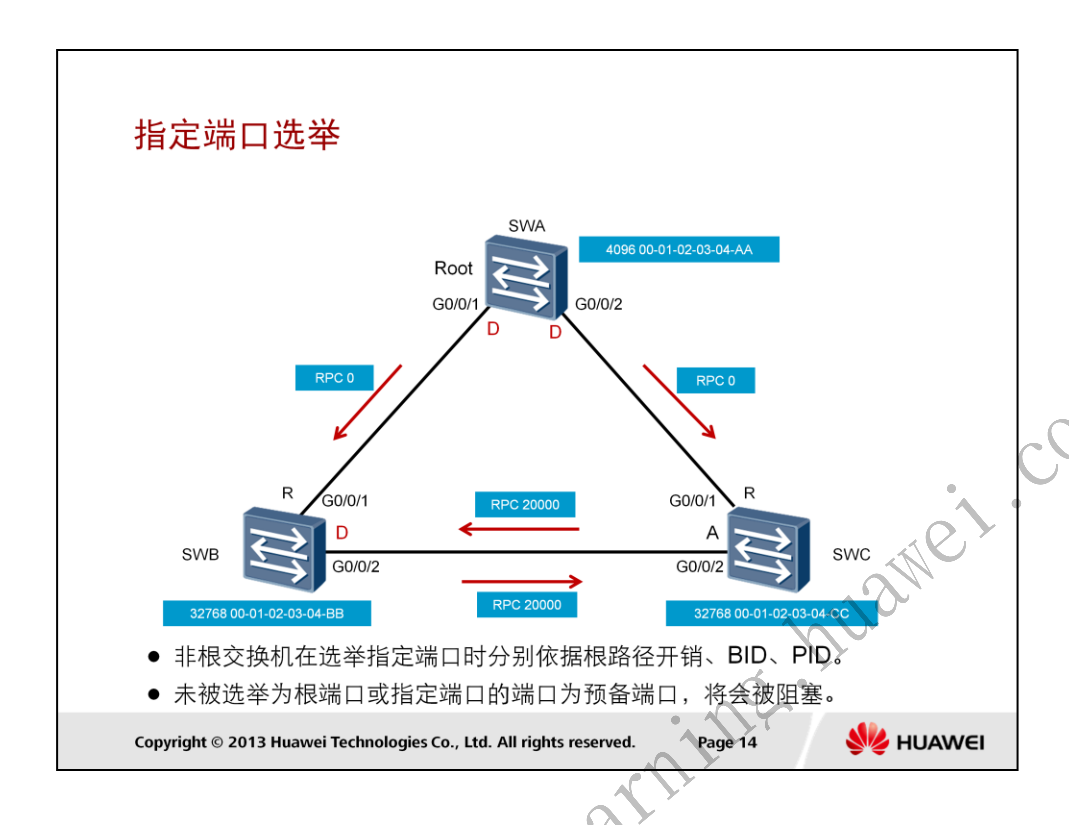
交换机的每个端口都有一个端口开销（Port Cost）参数，此参数表示该端口发送数据时的开销值，即出端口的开销。STP认为从一个端口接收数据是没有开销的。端口的开销和端口的带宽有关，带宽越高，开销越小。从一个非根桥到达根桥的路径可能有多条，每一条路径都有一个总的开销值，此开销值是该路径上所有出端口的端口开销总和，即根路径开销RPC（Root Path Cost）。非根桥根据根路径开销来确定到达根桥的最短路径，并生成无环树状网络。根桥的根路径开销是0。

一般情况下，企业网络中会存在多厂商的交换设备，华为X7系列交换机支持多种STP的路径开销计算标准，提供最大程度的兼容性。缺省情况下，华为X7系列交换机使用IEEE 802.1t标准来计算路径开销。

运行STP交换机的每个端口都有一个端口ID，端口ID由端口优先级和端口号构成。端口优先级取值范围是0到240，步长为16，即取值必须为16的整数倍。缺省情况下，端口优先级是128。端口ID（port ID)可以用来确定端口角色。

每个非根桥都要选举一个根端口。根端口是距离根桥最近的端口，这个最近的衡量标准是靠累计根路径开销来判定的，即累计根路径开销最小的端口就是根端口。端口收到一个BPDU报文后，抽取该BPDU报文中累计根路径开销字段的值，加上该端口本身的路径开销即为累计根路径开销。如果有两个或两个以上的端口计算得到的累计根路径开销相同，那么选择收到发送者BID最小的那个端口作为根端口。

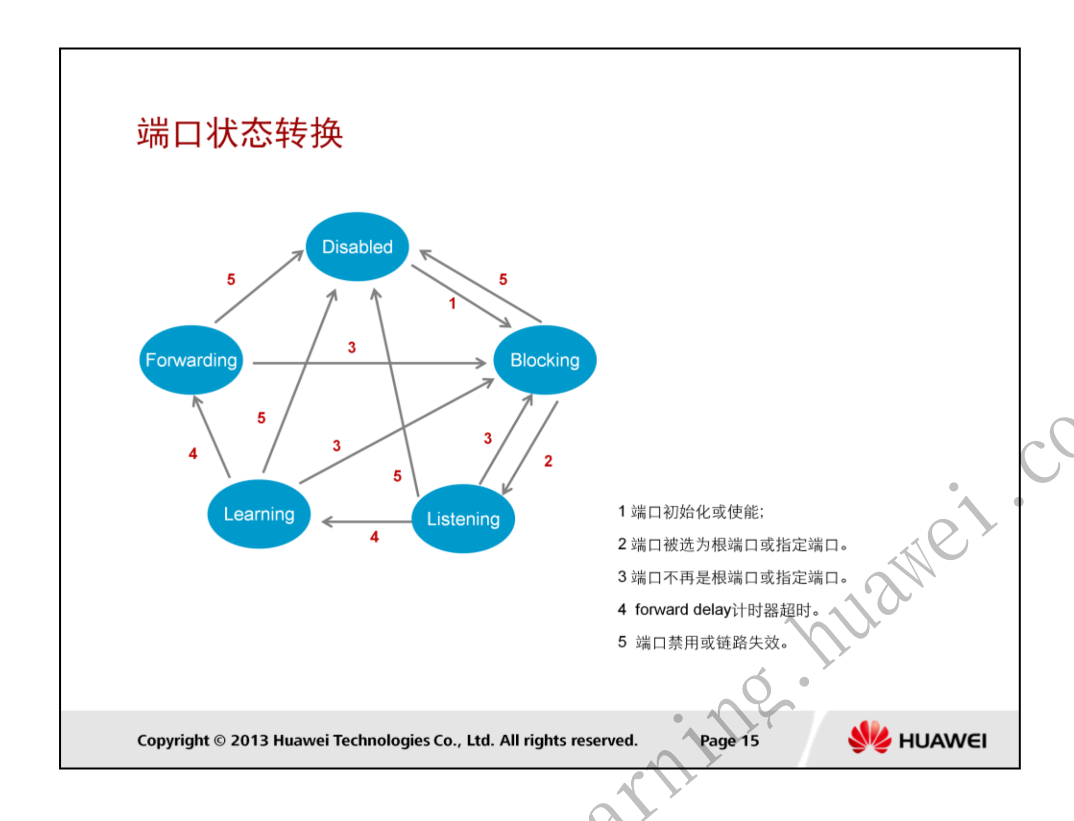
如果两个或两个以上的端口连接到同一台交换机上，则选择发送者PID最小的那个端口作为根端口。如果两个或两个以上的端口通过Hub连接到同一台交换机的同一个接口上，则选择本交换机的这些端口中的PID最小的作为根端口。



在网段上抑制其他端口（无论是自己的还是其他设备的）发送BPDU报文的端口，就是该网段的指定端口。每个网段都应该有一个指定端口，根桥的所有端口都是指定端口（除非根桥在物理上存在环路）。

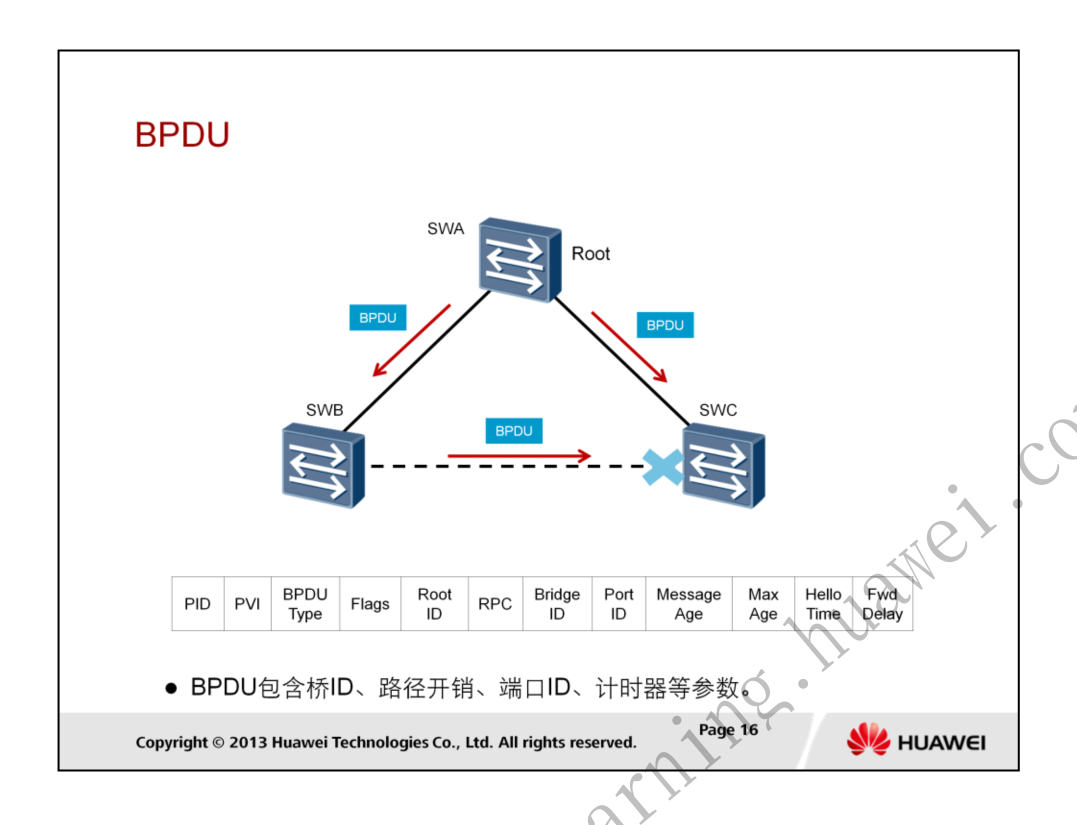
指定端口的选举也是首先比较累计根路径开销，累计根路径开销最小的端口就是指定端口。如果累计根路径开销相同，则比较端口所在交换机的桥ID，所在桥ID最小的端口被选举为指定端口。如果通过累计根路径开销和所在桥ID选举不出来，则比较端口ID，端口ID最小的被选举为指定端口。

网络收敛后，只有指定端口和根端口可以转发数据。其他端口为预备端口，被阻塞，不能转发数据，只能够从所连网段的指定交换机接收到BPDU报文，并以此来监视链路的状态。



图中所示为STP的端口状态迁移机制，运行STP协议的设备上端口状态有5种：

1. Forwarding：转发状态。端口既可转发用户流量也可转发BPDU报文，只有根端口或指定端口才能进入Forwarding状态。
2. Learning：学习状态。端口可根据收到的用户流量构建MAC地址表，但不转发用户流量。增加Learning状态是为了防止临时环路。
3. Listening：侦听状态。端口可以转发BPDU报文，但不能转发用户流量。
4. Blocking：阻塞状态。端口仅仅能接收并处理BPDU，不能转发BPDU，也不能转发用户流量。此状态是预备端口的最终状态。
5. Disabled：禁用状态。端口既不处理和转发BPDU报文，也不转发用户流量。



为了计算生成树，交换机之间需要交换相关的信息和参数，这些信息和参数被封装在BPDU（Bridge Protocol Data Unit）中。

BPDU有两种类型：配置BPDU和TCN BPDU。

1. 配置BPDU包含了桥ID、路径开销和端口ID等参数。STP协议通过在交换机之间传递配置BPDU来选举根交换机,以及确定每个交换机端口的角色和状态。在初始化过程中，每个桥都主动发送配置BPDU。在网络拓扑稳定以后，只有根桥主动发送配置BPDU，其他交换机在收到上游传来的配置BPDU后，才会发送自己的配置BPDU。
2. TCN BPDU是指下游交换机感知到拓扑发生变化时向上游发送的拓扑变化通知。

配置BPDU中包含了足够的信息来保证设备完成生成树计算，其中包含的重要信息如下：

根桥ID：由根桥的优先级和MAC地址组成，每个STP网络中有且仅有一个根。

根路径开销：到根桥的最短路径开销。

指定桥ID：由指定桥的优先级和MAC地址组成。

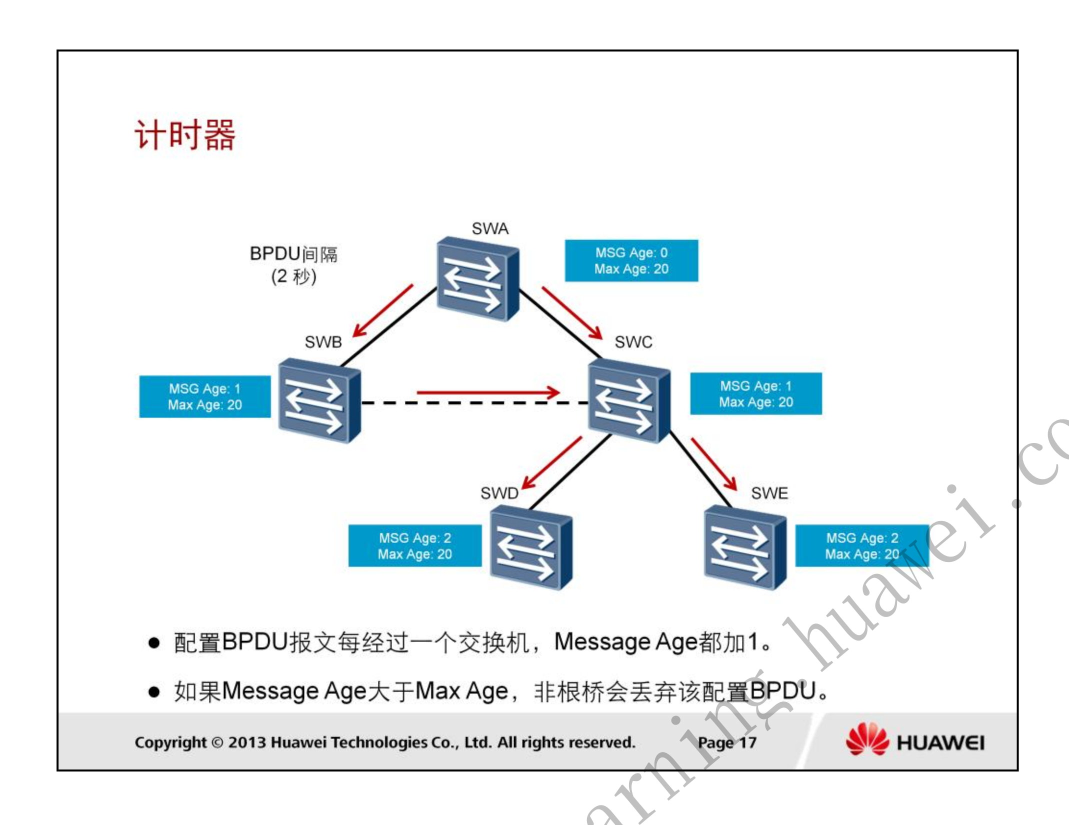
指定端口ID：由指定端口的优先级和端口号组成。

Message Age：配置BPDU在网络中传播的生存期。

Max Age：配置BPDU在设备中能够保存的最大生存期。

Hello Time：配置BPDU发送的周期。

Forward Delay：端口状态迁移的延时。

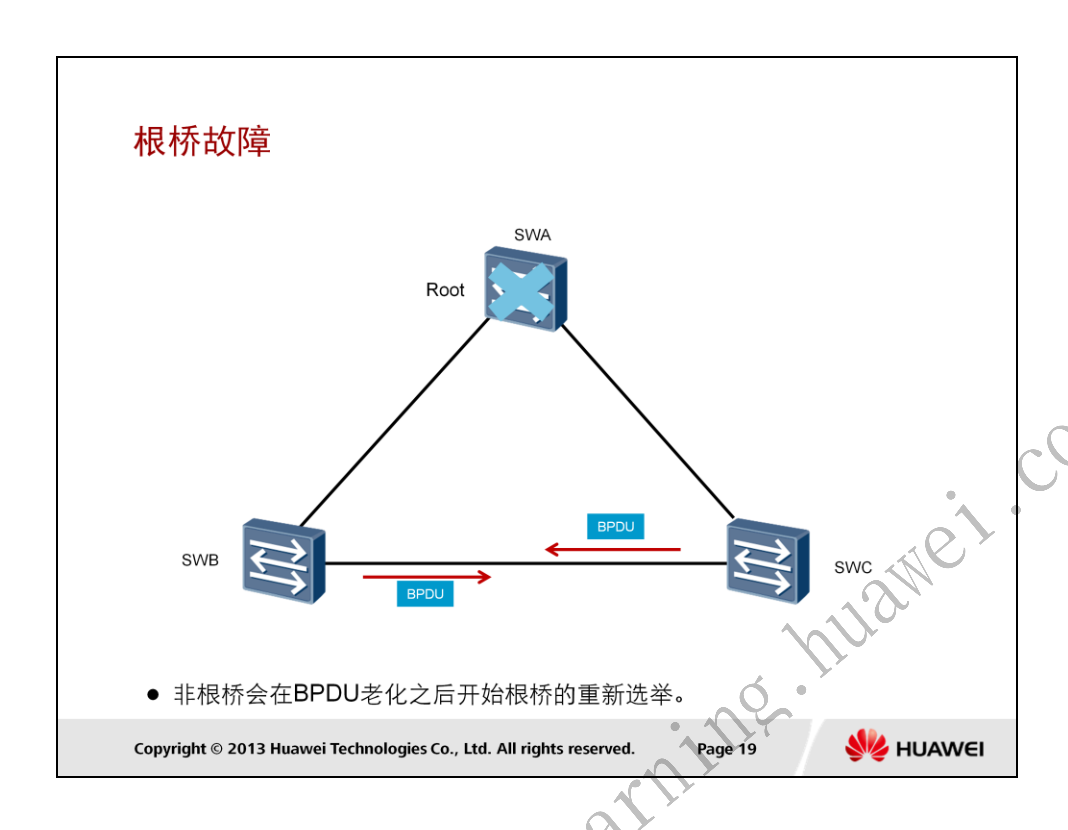


STP协议中包含一些重要的时间参数，这里举例说明如下：

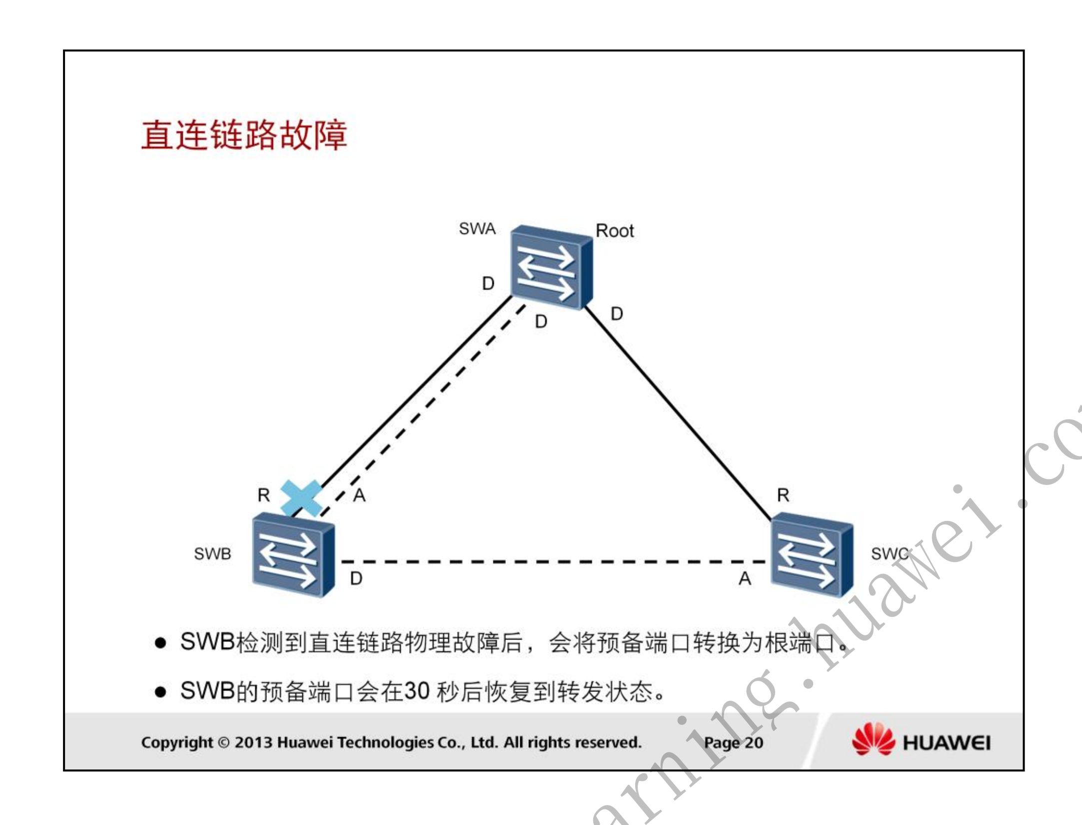
1. Hello Time是指运行STP协议的设备发送配置BPDU的时间间隔，用于检测链路是否存在故障。交换机每隔Hello Time时间会向周围的交换机发送配置BPDU报文，以确认链路是否存在故障。当网络拓扑稳定后，该值只有在根桥上修改才有效。
2. Message Age:如果配置BPDU是根桥发出的，则Message Age为0。否则，Message Age是从根桥发送到当前桥接收到BPDU的总时间，包括传输延时等。实际实现中，配置BPDU报文每经过一个交换机，Message Age增加1。
3. Max Age是指BPDU报文的老化时间，可在根桥上通过命令人为改动这个值。Max Age通过配置BPDU报文的传递，可以保证Max Age在整网中一致。非根桥设备收到配置BPDU报文后，会将报文中的Message Age和Max Age进行比较：如果Message Age小于等于Max Age，则该非根桥设备会继续转发配置BPDU报文。如果Message Age大于Max Age，则该配置BPDU报文将被老化掉。该非根桥设备将直接丢弃该配置BPDU，并认为是网络直径过大，导致了根桥连接失败。



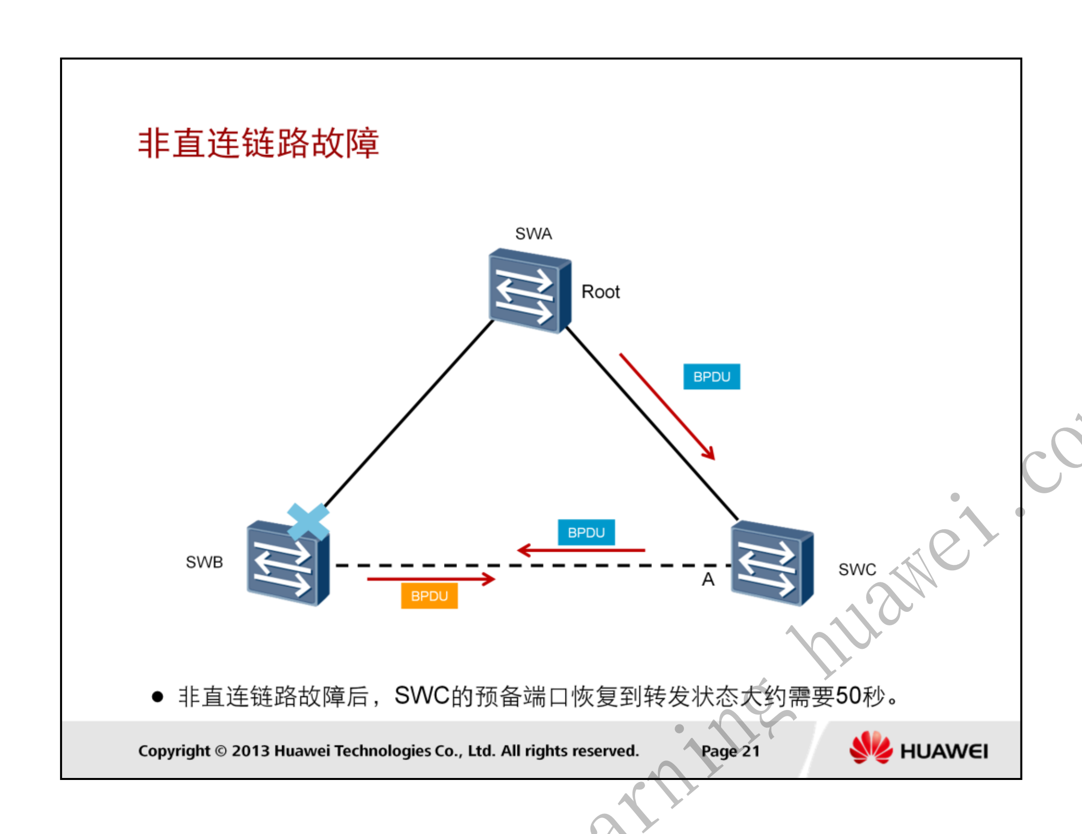
目录



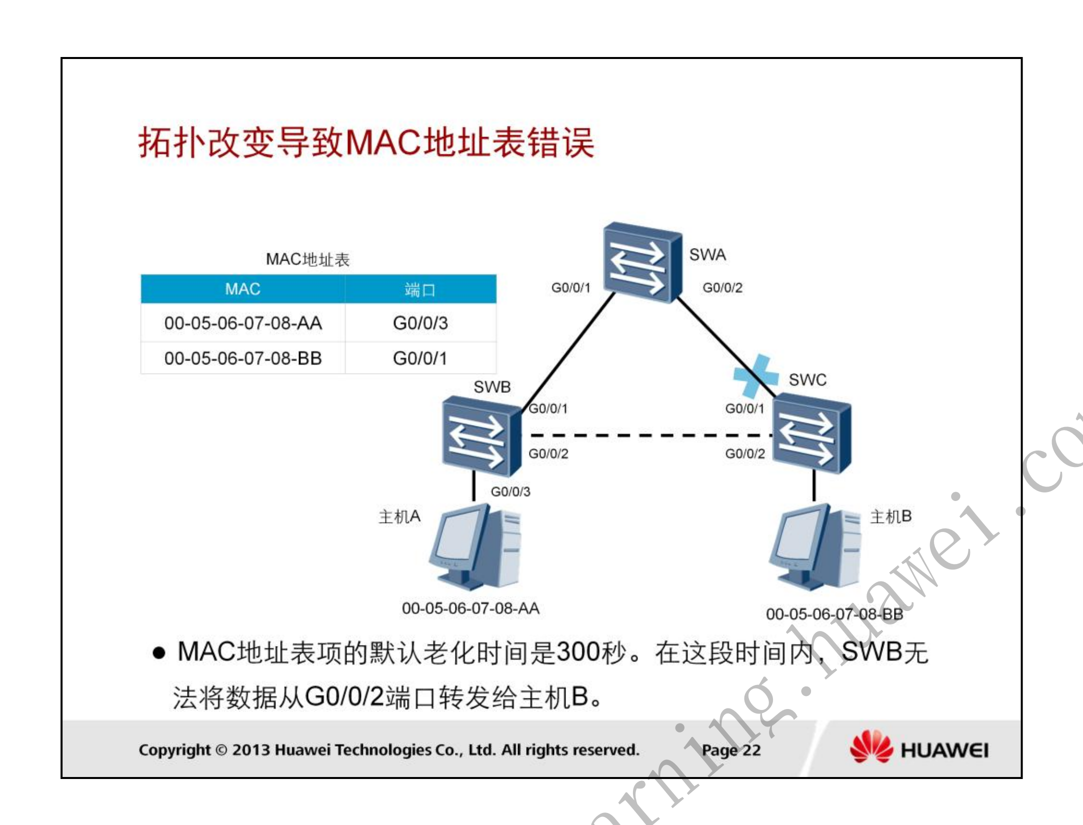
在稳定的STP拓扑里，非根桥会定期收到来自根桥的BPDU报文。如果根桥发生了故障，停止发送BPDU报文，下游交换机就无法收到来自根桥的BPDU报文。如果下游交换机一直收不到BPDU报文，Max Age定时器就会超时（Max Age的默认值为20秒），从而导致已经收到的BPDU报文失效，此时，非根交换机会互相发送配置BPDU报文，重新选举新的根桥。根桥故障会导致50秒左右的恢复时间，恢复时间约等于Max Age加上两倍的Forward Delay收敛时间。



此例中，SWA和SWB使用了两条链路互连，其中一条是主用链路，另外一条是备份链路。生成树正常收敛之后，如果SWB检测到根端口的链路发生物理故障，则其Alternate端口会迁移到Listening、 Learning、Forwarding状态，经过2倍的Forward Delay后恢复到转发状态。

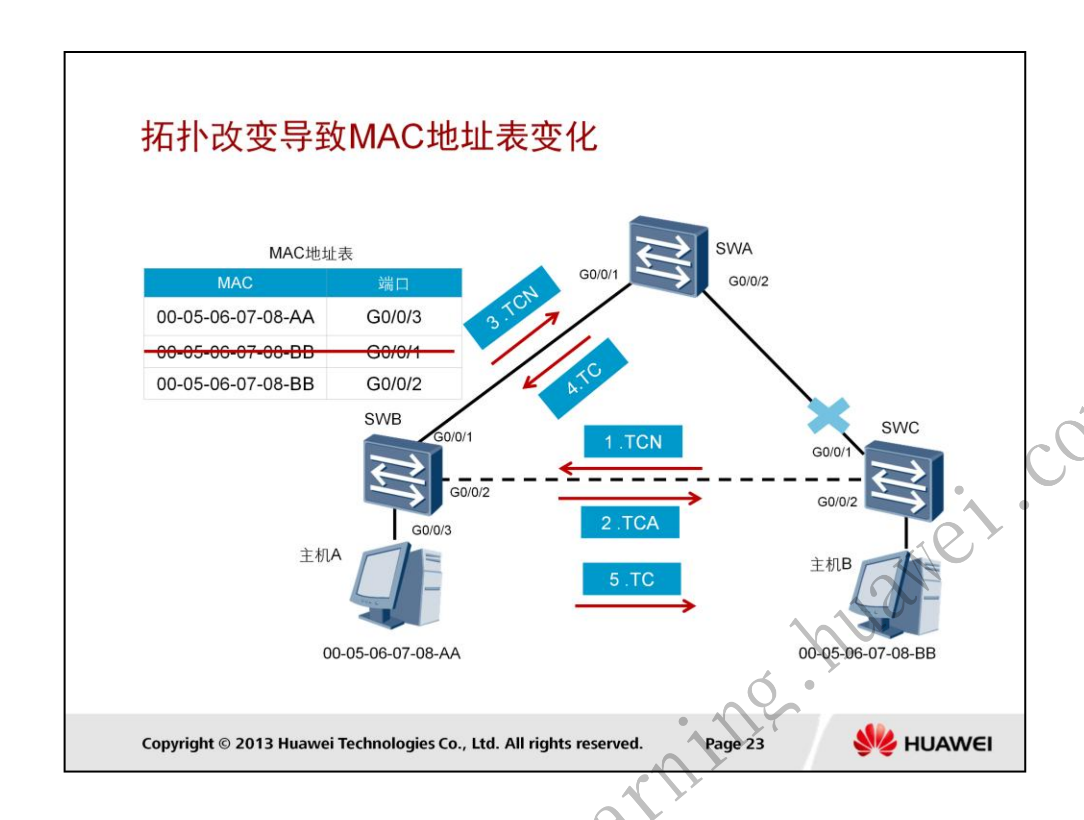


本例中，SWB与SWA之间的链路发生了某种故障（非物理层故障），SWB因此一直收不到来自SWA的BPDU报文。此时，SWB会认为根桥SWA不再有效，于是开始发送BPDU报文给SWC，通知SWC自己作为新的根桥。SWC也会继续从原根桥接收BPDU报文，因此会忽略SWB发送的BPDU报文。由于SWC的Alternate端口再也不能收到包含原根桥ID的BPDU报文。其Max Age定时器超时后，SWC会切换Alternate端口为指定端口并且转发来自其根端口的BPDU报文给SWB。SWB放弃宣称自己是根桥并开始收敛端口为根端口。非直连链路故障后，由于需要等待Max Age加上两倍的Forward Delay时间，端口需要大约50秒才能恢复到转发状态。



在交换网络中，交换机依赖MAC地址表转发数据帧。缺省情况下，MAC地址表项的老化时间是300秒。如果生成树拓扑发生变化，交换机转发数据的路径也会随着发生改变，此时MAC地址表中未及时老化掉的表项会导致数据转发错误，因此在拓扑发生变化后需要及时更新MAC地址表项。

本例中，SWB中的MAC地址表项定义了通过端口GigabitEthernet 0/0/3可以到达主机A，通过端口GigabitEthernet 0/0/1可以到达主机B。由于SWC的根端口产生故障，导致生成树拓扑重新收敛，在生成树拓扑完成收敛之后，从主机A到主机B的帧仍然不能到达目的地。这是因为MAC地址表项老化时间是300秒，主机A发往主机B的帧到达SWB后，SWB会继续通过端口GigabitEthernet 0/0/1转发该数据帧。



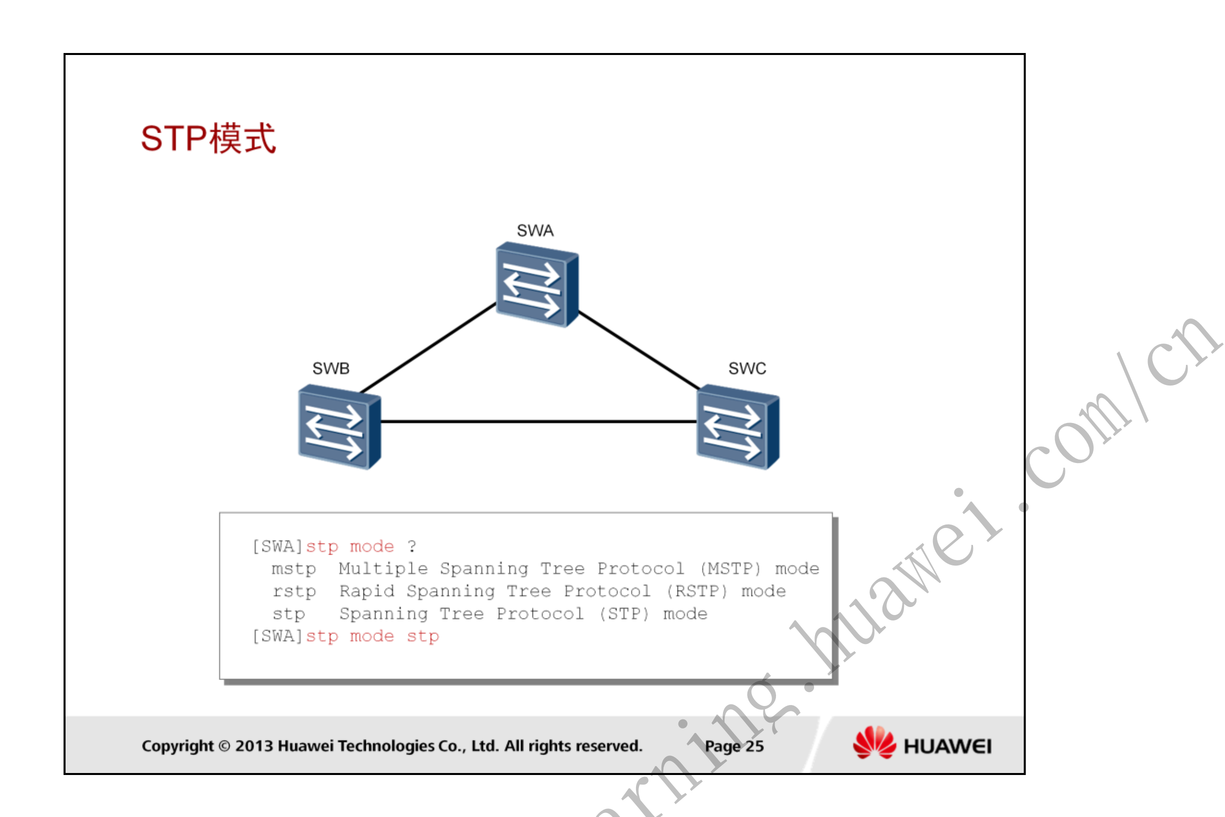
拓扑变化过程中，根桥通过TCN BPDU报文获知生成树拓扑里发生了故障。根桥生成TC用来通知其他交换机加速老化现有的MAC地址表项。

拓扑变更以及MAC地址表项更新的具体过程如下：

1. SWC感知到网络拓扑发生变化后，会不间断地向SWB发送TCN BPDU报文。
2. SWB收到SWC发来的TCN BPDU报文后，会把配置BPDU报文中的Flags的TCA位设置1，然后发送给SWC，告知SWC停止发送TCN BPDU报文。
3. SWB向根桥转发TCN BPDU报文。
4. SWA把配置BPDU报文中的Flags的TC位设置为1后发送，通知下游设备把MAC地址表项的老化时间由默认的300秒修改为Forwarding Delay的时间（默认为15秒）。
5. 最多等待15秒之后，SWB中的错误映射关系会被自动清除。此后，SWB就能通过G0/0/2端口把从主机A到主机B的帧正确地进行转发。

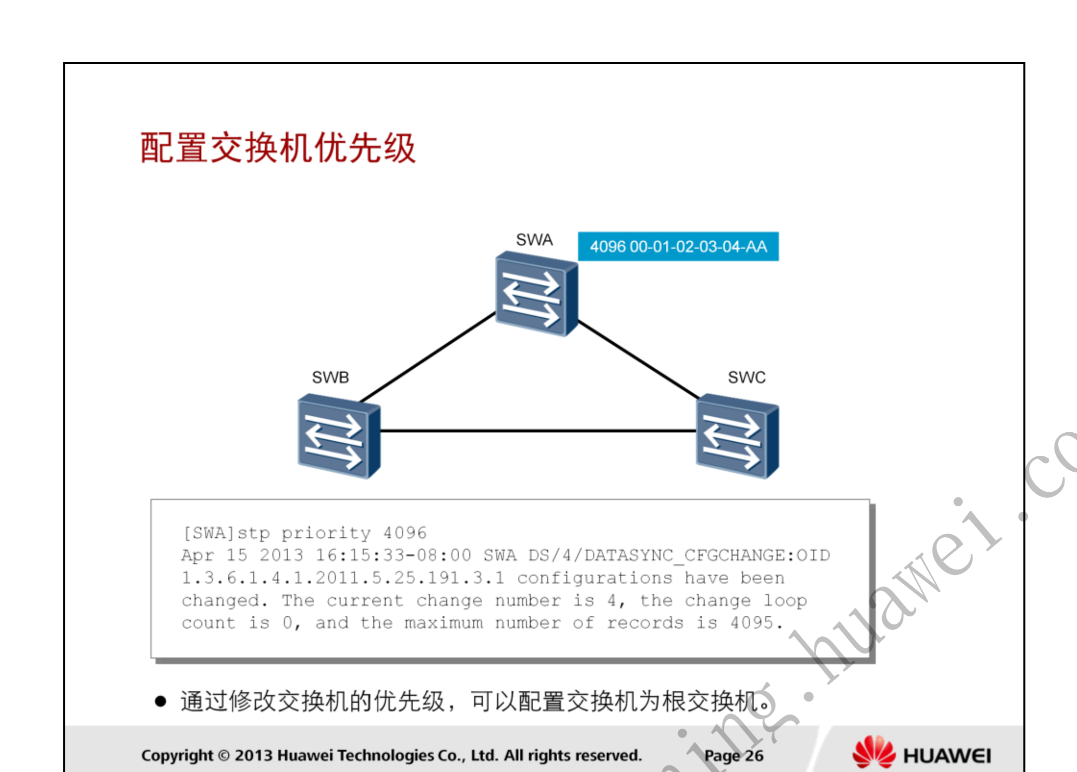


目录



华为X7系列交换机支持三种生成树协议模式。

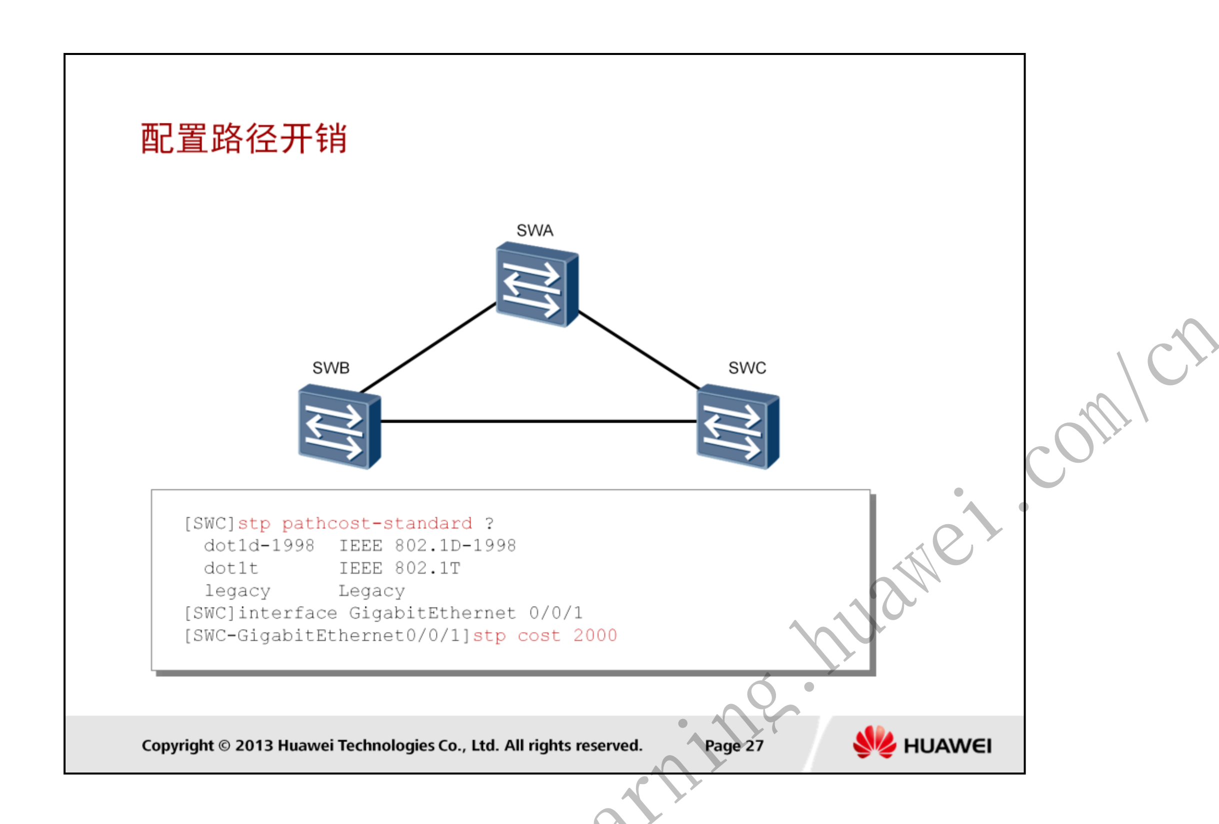
stp mode { **mstp** | **stp** | **rstp** }命令用来配置交换机的生成树协议模式。缺省情况下，华为X7系列交换机工作在MSTP模式。在使用STP前，STP模式必须重新配置。



基于企业业务对网络的需求，一般建议手动指定网络中配置高、性能好的交换机为根桥。

可以通过配置桥优先级来指定网络中的根桥，以确保企业网络里面的数据流量使用最优路径转发。

stp priority *priority*命令用来配置设备优先级值。*priority*值为整数，取值范围为0到61440，步长为4096。缺省情况下，交换设备的优先级取值是32768。另外，可以通过**stp root primary**命令指定生成树里的根桥。



华为X7系列交换机支持三种路径开销标准，以确保和友商设备保持兼容。缺省情况下，路径开销标准为IEEE 802.1t。

stp pathcost-standard { dot1d-1998 | dot1t | legacy }命令用来配置指定交换机上路径开销值的标准。

每个端口的路径开销也可以手动指定。此STP路径开销控制方法须谨慎使用，手动指定端口的路径开销可能会生成次优生成树拓扑。

stp cost *cost*命令取决于路径开销计算方法：

1. 使用华为的私有计算方法时，*cost*取值范围是1～200000。
2. 使用IEEE 802.1d标准方法时，*cost*取值范围是1～65535。
3. 使用IEEE 802.1t标准方法时，*cost*取值范围是1～200000000。

配置验证

```
[SWA]display stp
-------[CIST Global Info][Mode STP]-------
CIST Bridge          :4096 .00-01-02-03-04-BB
Bridge Times         :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
CIST Root/ERPC       :4096 .00-01-02-03-04-BB / 0
CIST RegRoot/IRPC    :4096 .00-01-02-03-04-BB / 0
CIST RootPortId      :0.0
BPDU-Protection      :Disabled
TC or TCN received   :37
TC count per hello   :0
STP Converge Mode    :Normal
Share region-configuration :Enabled
Time since last TC   :0 days 0h:1m:29s
……
```

 Page 28 HUAWEI

display stp命令用来检查当前交换机的STP配置。命令输出中信息介绍如下。

CIST Bridge参数标识指定交换机当前桥ID，包含交换机的优先级和MAC地址。

Bridge Times参数标识Hello定时器、Forward Delay定时器、Max Age定时器的值。

CIST Root/ERPC参数标识根桥ID以及此交换机到根桥的根路径开销。

配置验证

```
[SWA]display stp
……
----[Port1(GigabitEthernet0/0/1)][FORWARDING]----
 Port Protocol              :Enabled
 Port Role                  :Designated Port
 Port Priority              :128
 Port Cost(Dot1T )          :Config=2000 / Active=2000
 Designated Bridge/Port     :4096.00-01-02-03-04-BB / 128.1
 Port Edged                 :Config=default / Active=disabled
 Point-to-point             :Config=auto / Active=true
 Transit Limit              :147 packets/hello-time
 Protection Type            :None
 ……
```

 Page 29 HUAWEI

display stp命令显示交换机上所有端口信息；**display stp interface** *interface*命令显示交换机上指定端口信息。其他一些信息还包括端口角色、端口状态、以及使用的保护机制等。



总结

- 根桥产生故障后，其他交换机会被选举为根桥。那么原来的根桥恢复正常之后，网络又会发生什么变化呢？
- 路径开销和根路径开销的区别是什么？

 Page 30  

1. 如果生成树网络里面根桥发生了故障，则其它交换机中优先级最高的交换机会被选举为新的根桥。如果原来根桥再次激活，则网络又会根据BID来重新选举新的根桥。
2. 根路径开销是到根桥的路径的总开销，而路径开销指的是交换机端口的开销。


谢谢
www.huawei.com


RSTP原理与配置
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

STP协议虽然能够解决环路问题，但是收敛速度慢，影响了用户通信质量。如果STP网络的拓扑结构频繁变化，网络也会频繁失去连通性，从而导致用户通信频繁中断。IEEE于2001年发布的802.1w标准定义了快速生成树协议RSTP（Rapid Spanning-Tree Protocol），RSTP在STP基础上进行了改进，实现了网络拓扑快速收敛。

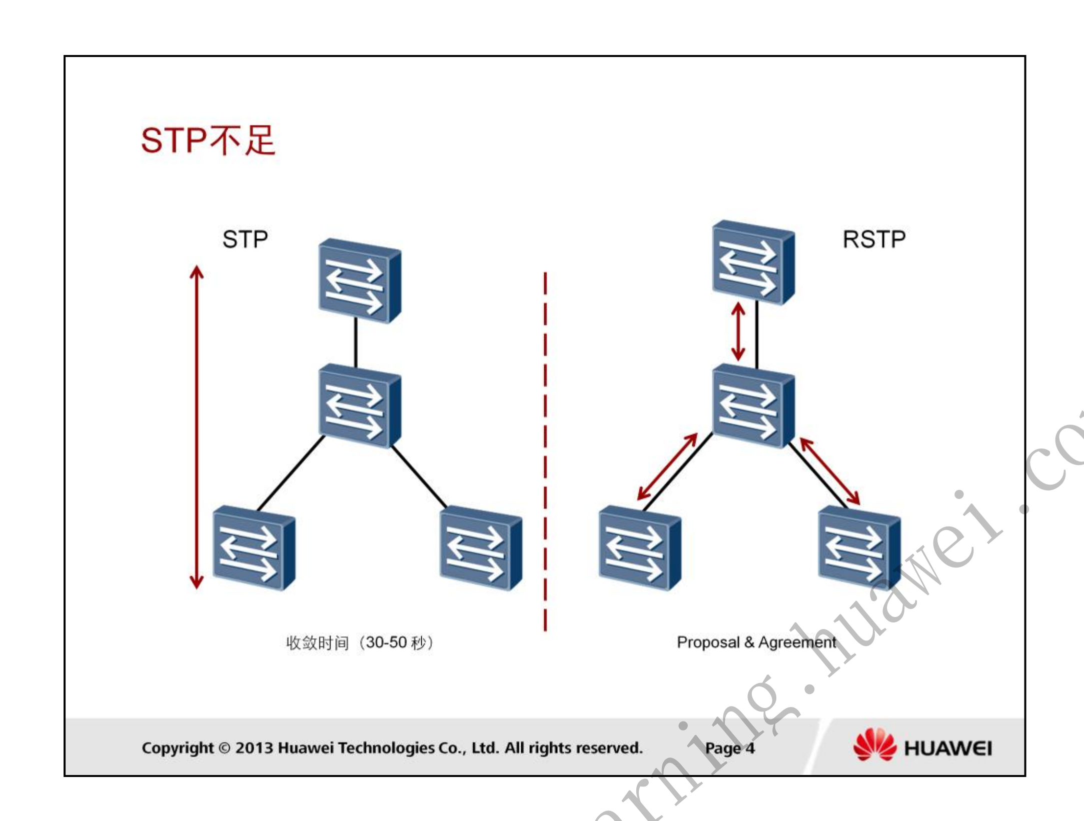




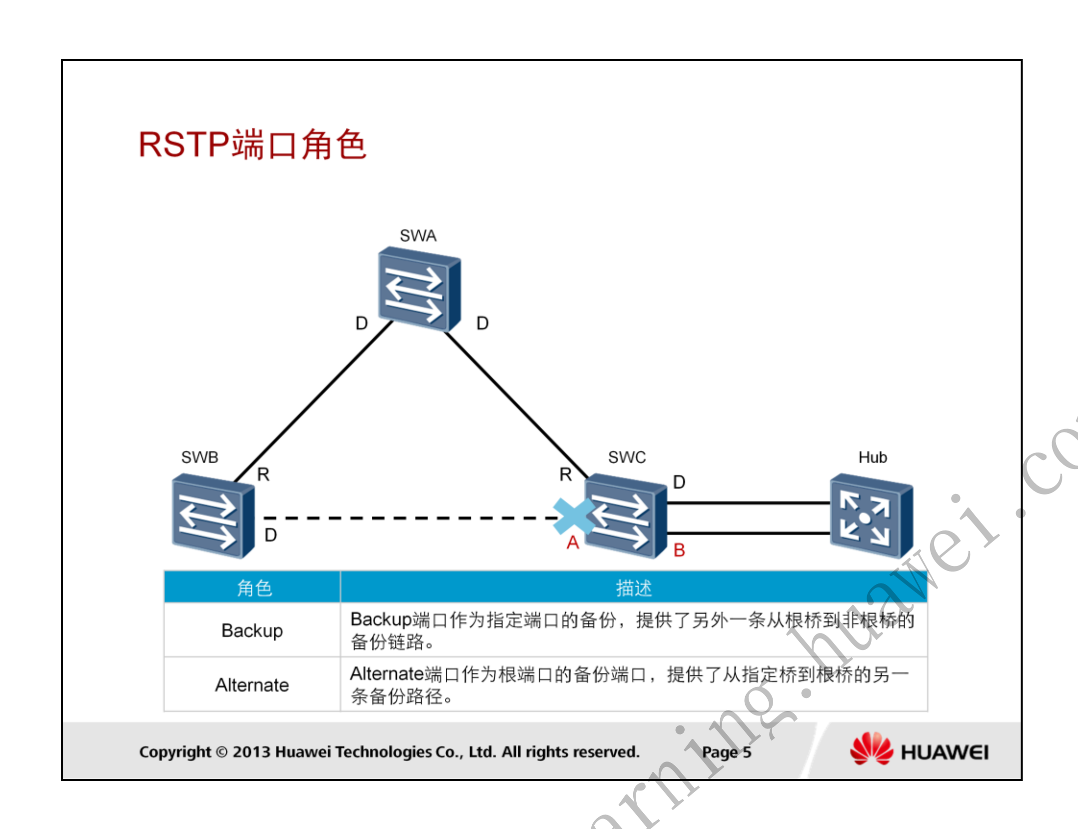
学习目标

学完本课程后，您应该能：

- 掌握RSTP的特性
- 配置RSTP

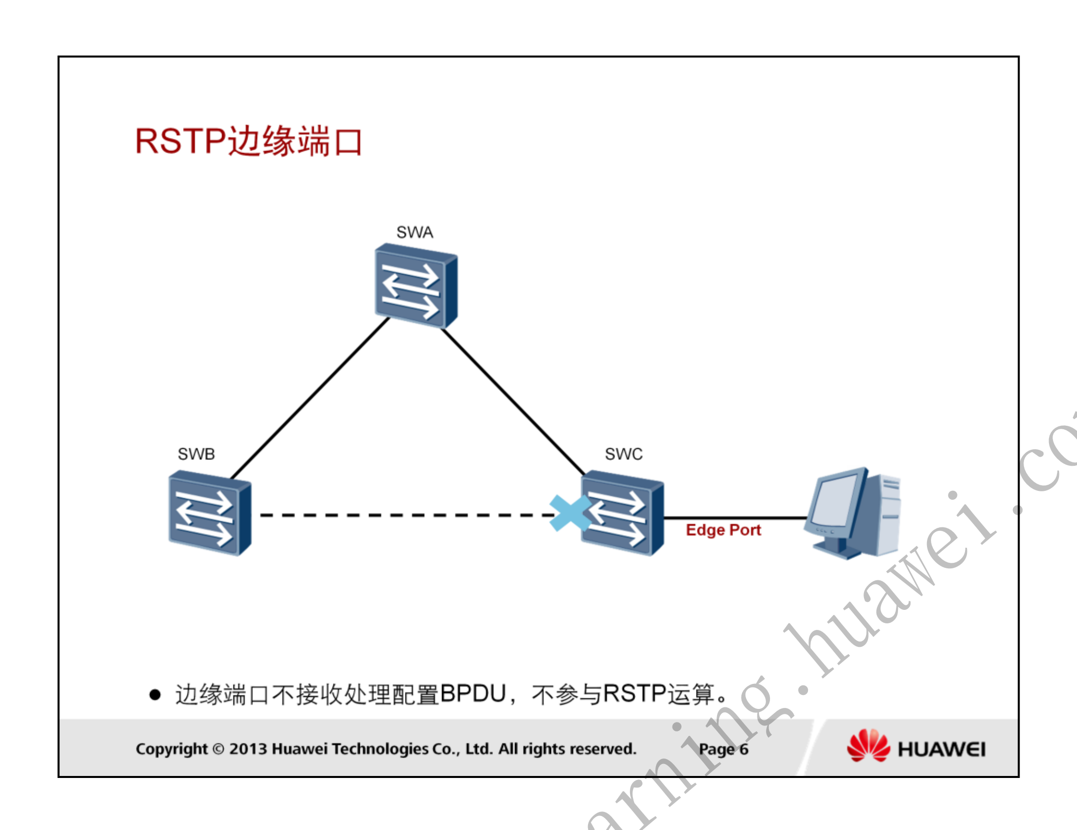


STP能够提供无环网络，但是收敛速度较慢。如果STP网络的拓扑结构频繁变化，网络也会随之频繁失去连通性，从而导致用户通信频繁中断。RSTP使用了Proposal/Agreement机制保证链路及时协商，从而有效避免收敛计时器在生成树收敛前超时。如图所示，在交换网络中，P/A过程可以从根桥向下游级联传递。



角色	描述
Backup	Backup端口作为指定端口的备份，提供了另外一条从根桥到非根桥的备份链路。
Alternate	Alternate端口作为根端口的备份端口，提供了从指定桥到根桥的另一条备份路径。

运行RSTP的交换机使用了两个不同的端口角色来实现冗余备份。当到根桥的当前路径出现故障时，作为根端口的备份端口，Alternate端口提供了从一个交换机到根桥的另一条可切换路径。Backup端口作为指定端口的备份，提供了另一条从根桥到相应LAN网段的备份路径。当一个交换机和一个共享媒介设备例如Hub建立两个或者多个连接时，可以使用Backup端口。同样，当交换机上两个或者多个端口和同一个LAN网段连接时，也可以使用Backup端口。



RSTP里，位于网络边缘的指定端口被称为边缘端口。边缘端口一般与用户终端设备直接连接，不与任何交换设备连接。边缘端口不接收配置BPDU报文，不参与RSTP运算，可以由Disabled状态直接转到Forwarding状态，且不经历时延，就像在端口上将STP禁用了一样。但是，一旦边缘端口收到配置BPDU报文，就丧失了边缘端口属性，成为普通STP端口，并重新进行生成树计算，从而引起网络震荡。

端口状态

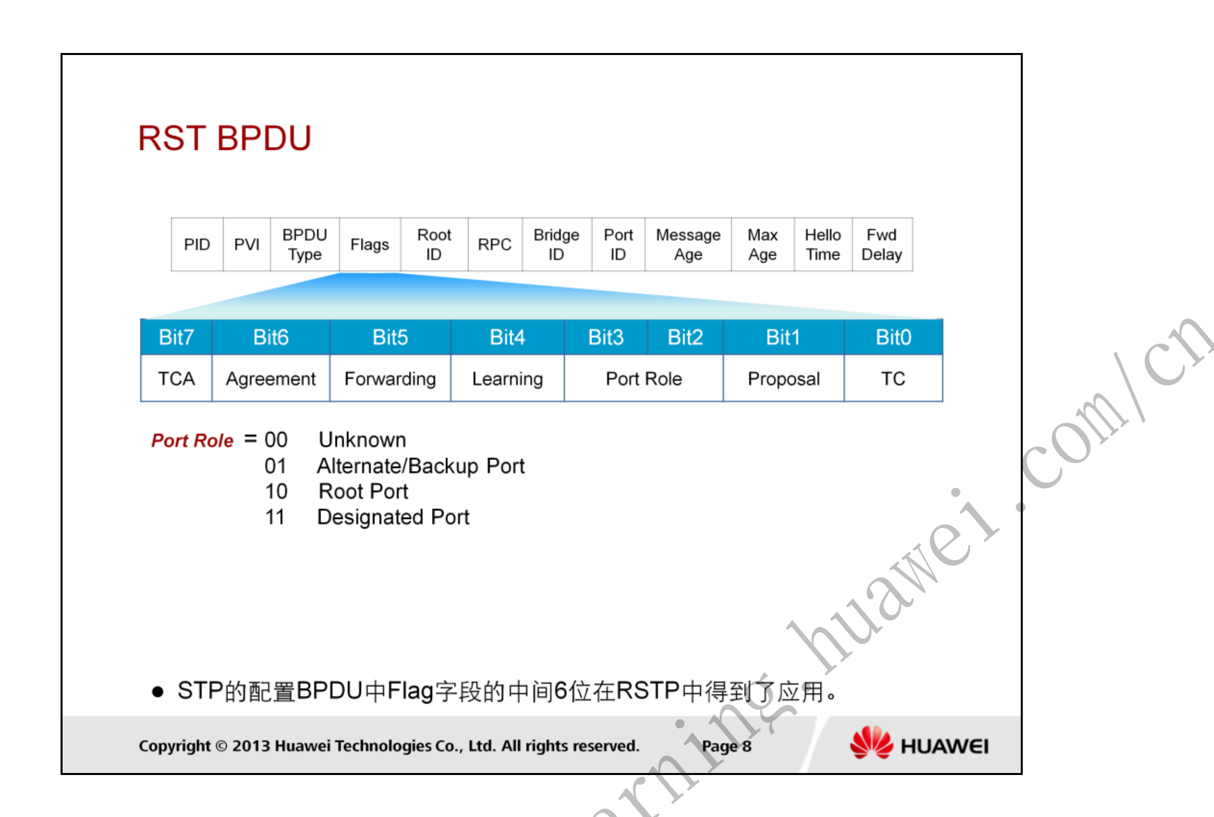
STP	RSTP	端口角色
Disabled	Discarding	Disable
Blocking	Discarding	Alternate端口、 Backup端口
Listening	Discarding	根端口、指定端口
Learning	Learning	根端口、指定端口
Forwarding	Forwarding	根端口、指定端口

 Page 7

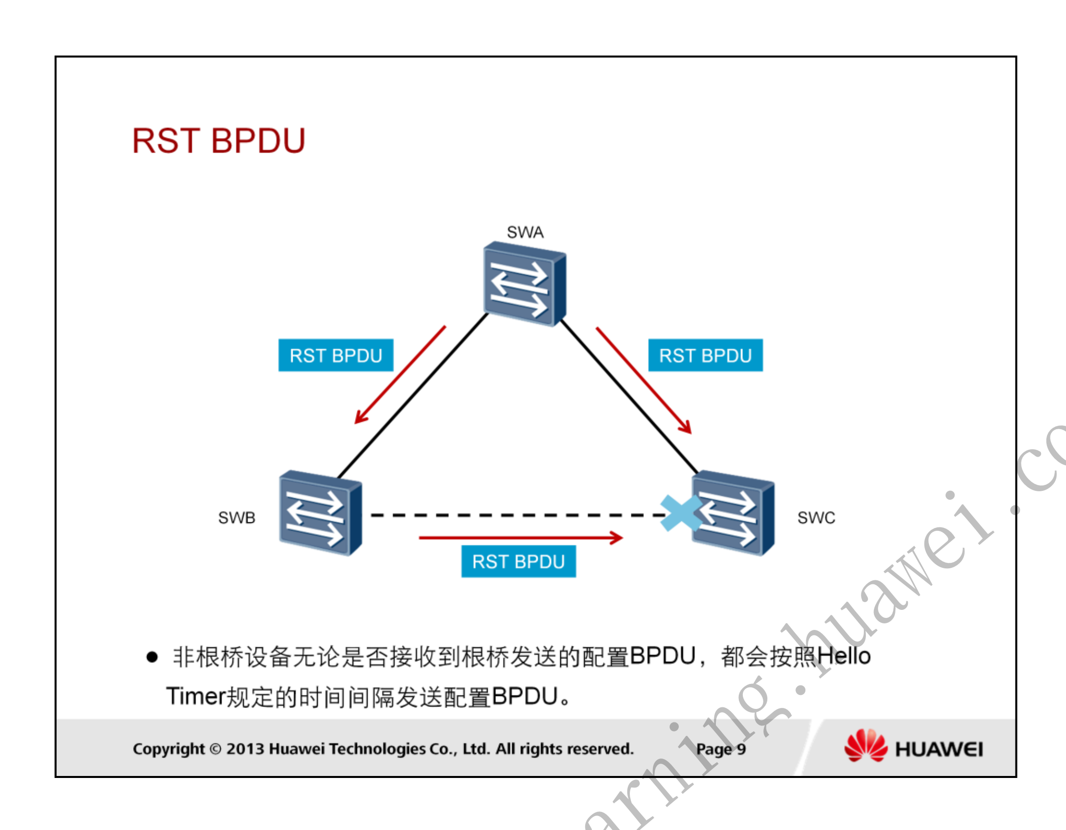


RSTP把原来STP的5种端口状态简化成了3种。

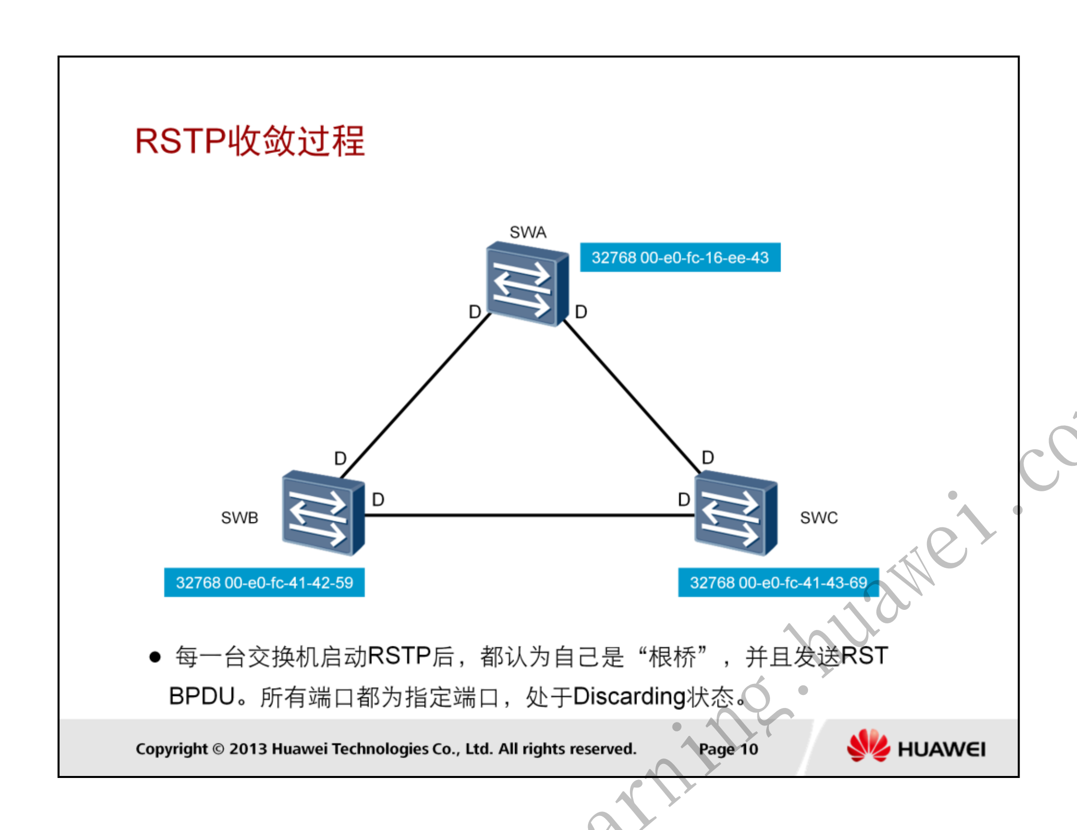
1. Discarding状态，端口既不转发用户流量也不学习MAC地址。
2. Learning状态，端口不转发用户流量但是学习MAC地址。
3. Forwarding状态，端口既转发用户流量又学习MAC地址。



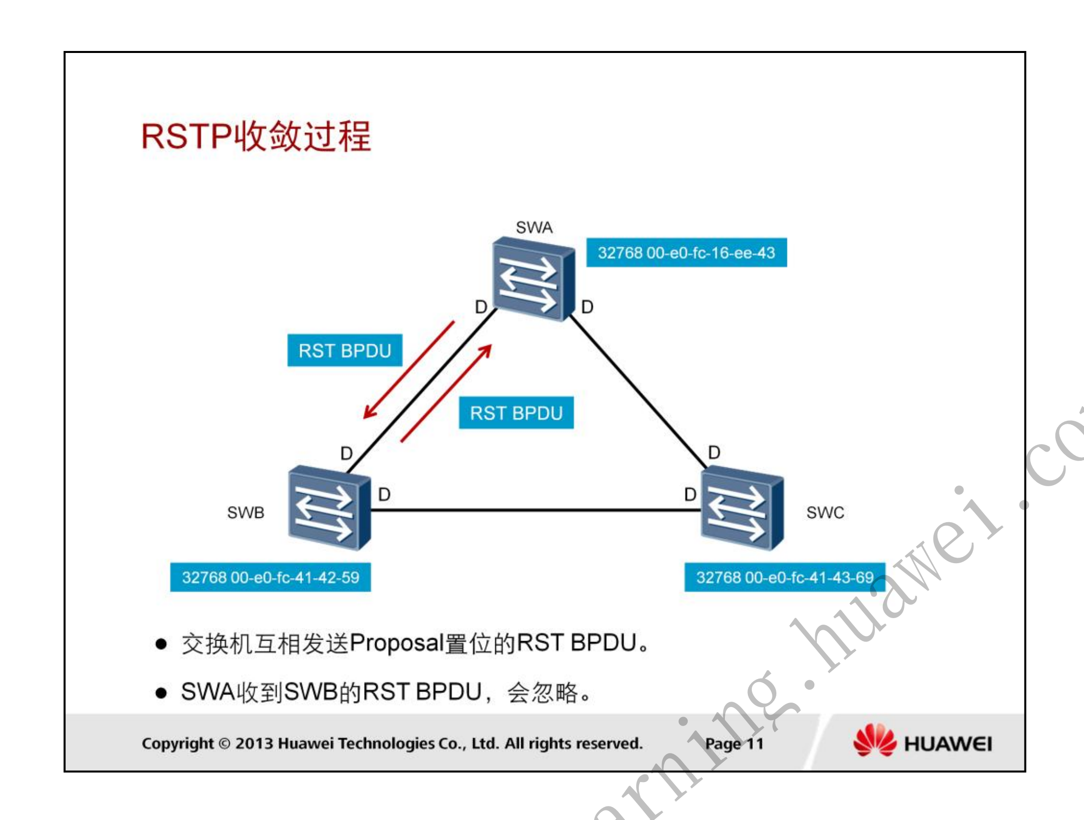
除了部分参数不同，RSTP使用了类似STP的BPDU报文，即RST BPDU报文。BPDU Type用来区分STP的BPDU报文和RST (Rapid Spanning Tree) BPDU报文。STP的配置BPDU报文的BPDU Type值为0(0x00)，TCN BPDU报文的BPDU Type值为128 (0x80），RST BPDU报文的BPDU Type值为2 (0x02)。STP的BPDU报文的Flags字段中只定义了拓扑变化TC（Topology Change）标志和拓扑变化确认TCA（Topology Change Acknowledgment）标志，其他字段保留。在RST BPDU报文的Flags字段里，还使用了其他字段。包括P/A进程字段和定义端口角色以及端口状态的字段。Forwarding，Learning与Port Role表示发出BPDU的端口的状态和角色。



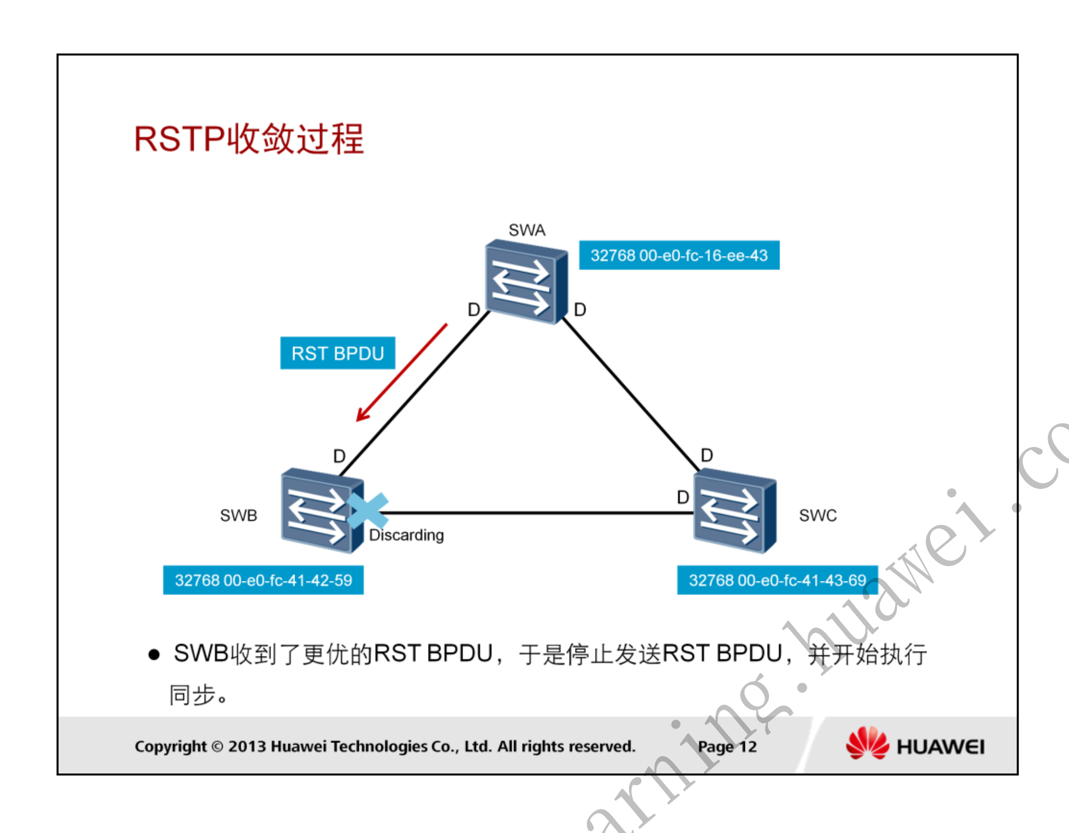
STP中，当网络拓扑稳定后，根桥按照Hello Timer规定的时间间隔发送配置BPDU报文，其他非根桥设备在收到上游设备发送过来的配置BPDU报文后，才会触发发出配置BPDU报文，此方式使得STP协议计算复杂且缓慢。RSTP对此进行了改进，即在拓扑稳定后，无论非根桥设备是否接收到根桥传来的配置BPDU报文，非根桥设备都会仍然按照Hello Timer规定的时间间隔发送配置BPDU，该行为完全由每台设备自主进行。



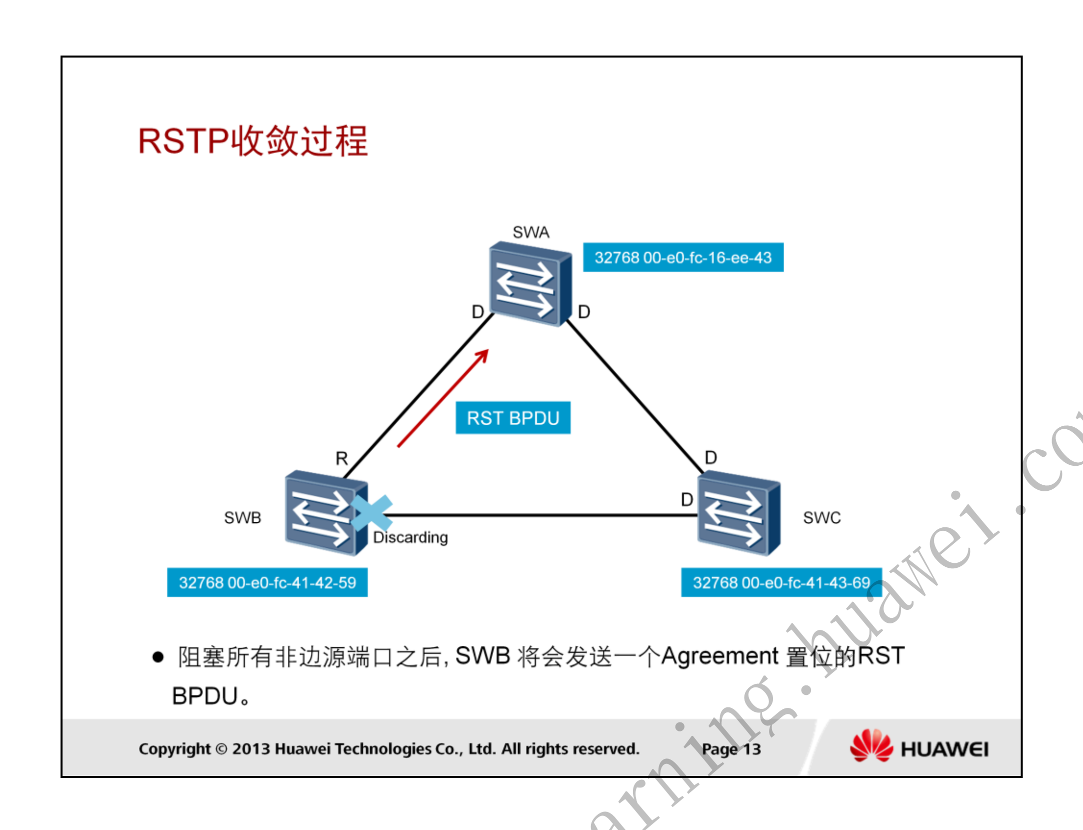
RSTP收敛遵循STP基本原理。网络初始化时，网络中所有的RSTP交换机都认为自己是“根桥”，并设置每个端口为指定端口。此时，端口为Discarding状态。



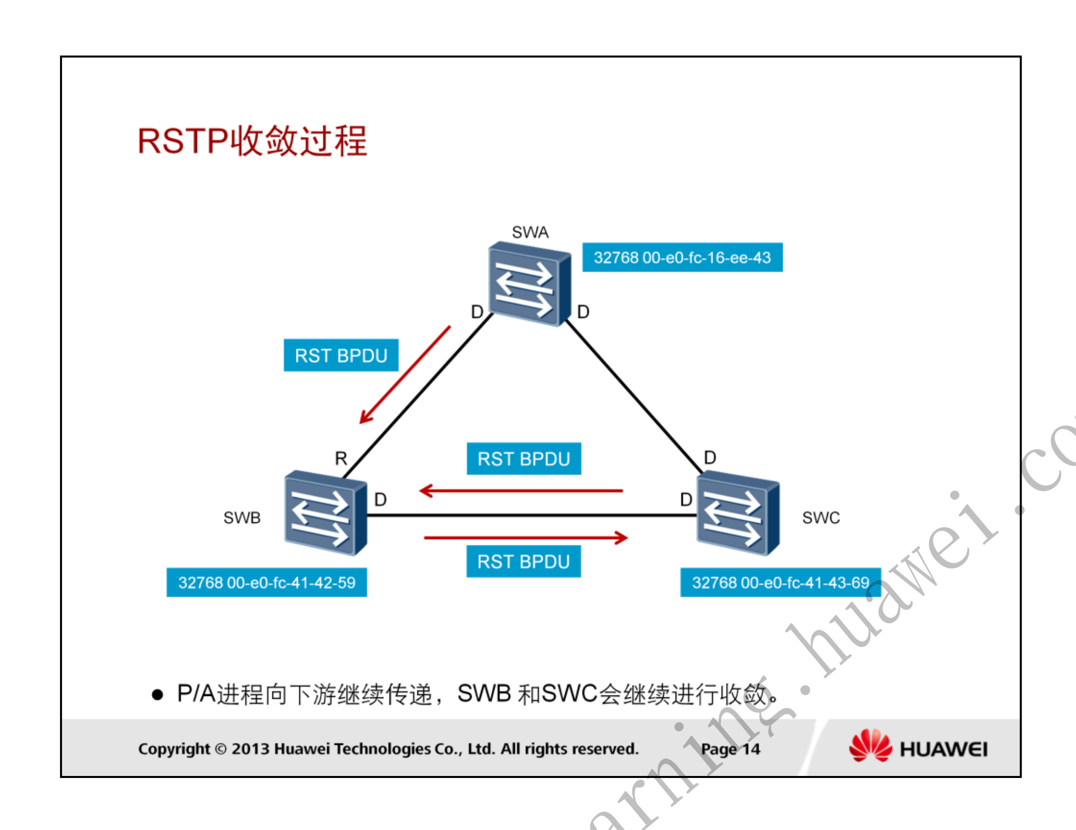
每个认为自己是“根桥”的交换机生成一个RST BPDU报文来协商指定网段的端口状态，此RST BPDU报文的Flags字段里面的Proposal位需要置位。当一个端口收到RST BPDU报文时，此端口会比较收到的RST BPDU报文和本地的RST BPDU报文。如果本地的RST BPDU报文优于接收的RST BPDU报文，则端口会丢弃接收的RST BPDU报文，并发送Proposal置位的本地RST BPDU报文来回复对端设备。



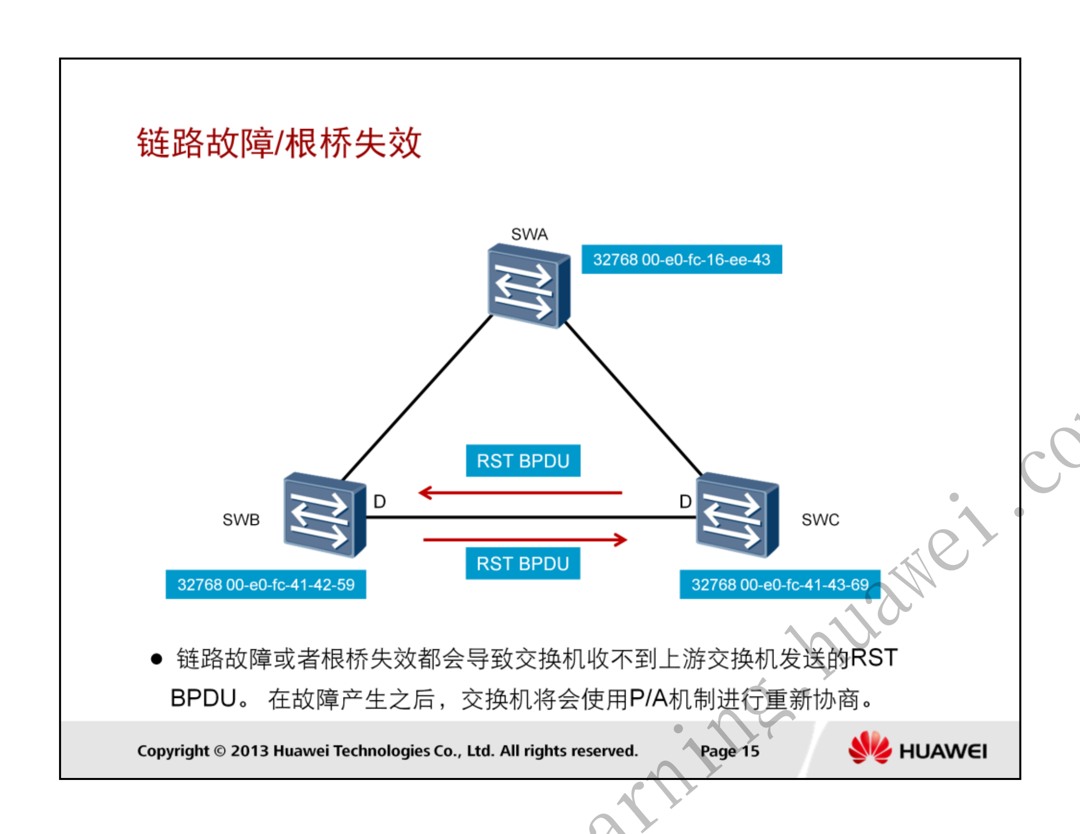
交换机使用同步机制来实现端口角色协商管理。当收到Proposal置位并且优先级高的BPDU报文时，接收交换机必须设置所有下游指定端口为Discarding状态。如果下游端口是Alternate端口或者边缘端口，则端口状态保持不变。本例说明了下游指定端口暂时迁移到Discarding状态的情形，因此，P/A进程中任何帧转发都将被阻止。



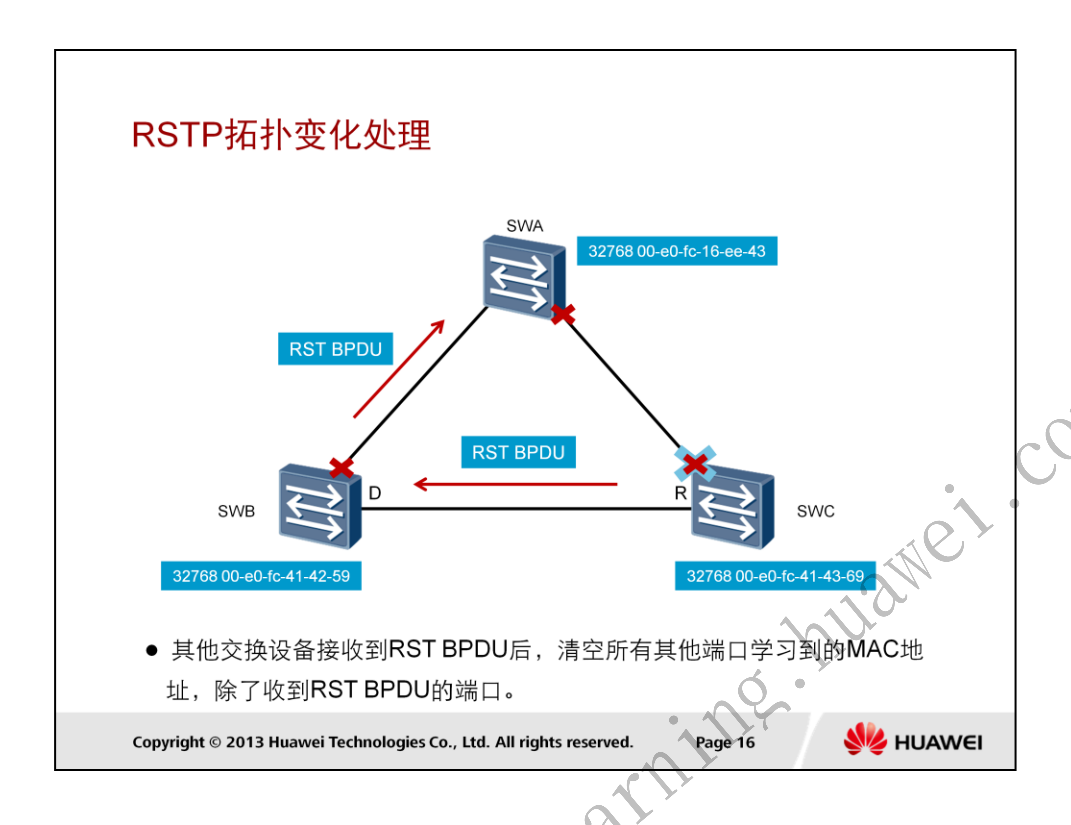
当确认下游指定端口迁移到Discarding状态后，设备发送RST BPDU报文回复上游交换机发送的Proposal消息。在此过程中，端口已经确认为根端口，因此RST BPDU报文Flags字段里面设置了Agreement标记位和根端口角色。



在P/A进程的最后阶段，上游交换机收到Agreement置位的RST BPDU报文后，指定端口立即从Discarding状态迁移为Forwarding状态。然后，下游网段开始使用同样的P/A进程协商端口角色。

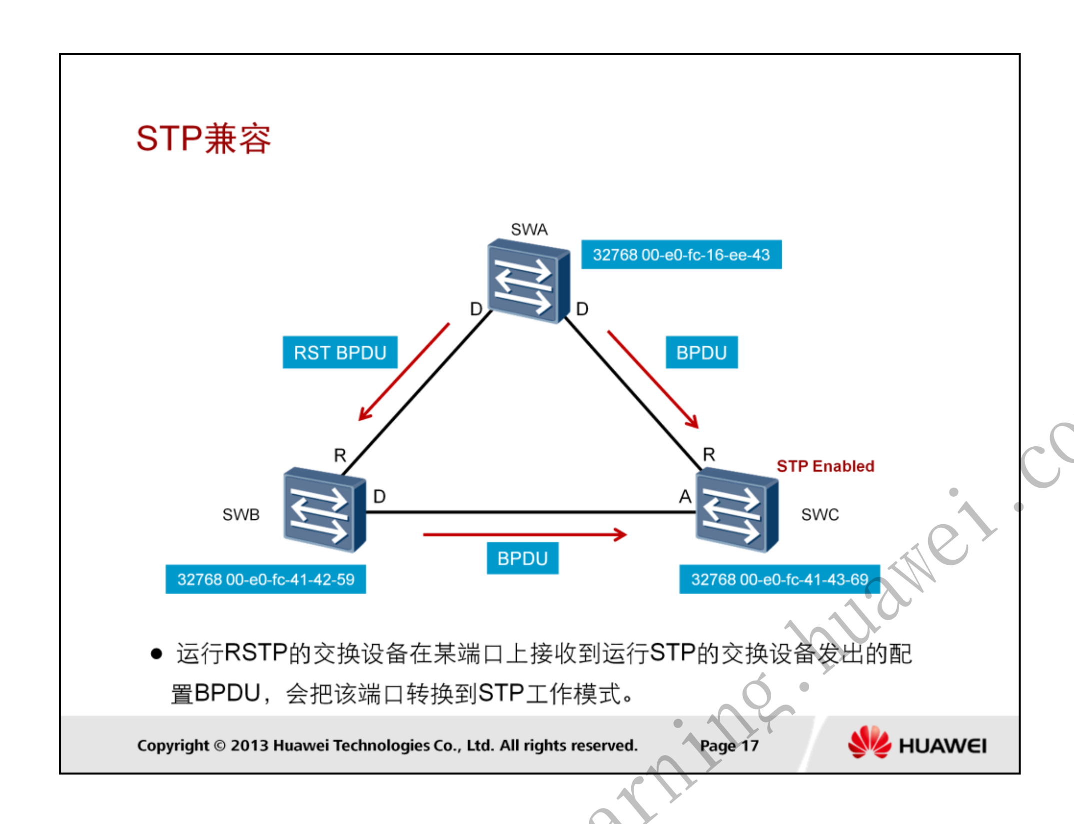


在STP中，当出现链路故障或根桥失效导致交换机收不到BPDU时，交换机需要等待Max Age时间后才能确认出现了故障。而在RSTP中，如果交换机的端口在连续3次Hello Timer规定的时间间隔内没有收到上游交换机发送的RST BPDU，便会确认本端口和对端端口的通信失败，从而需要初始化P/A进程去重新调整端口角色。



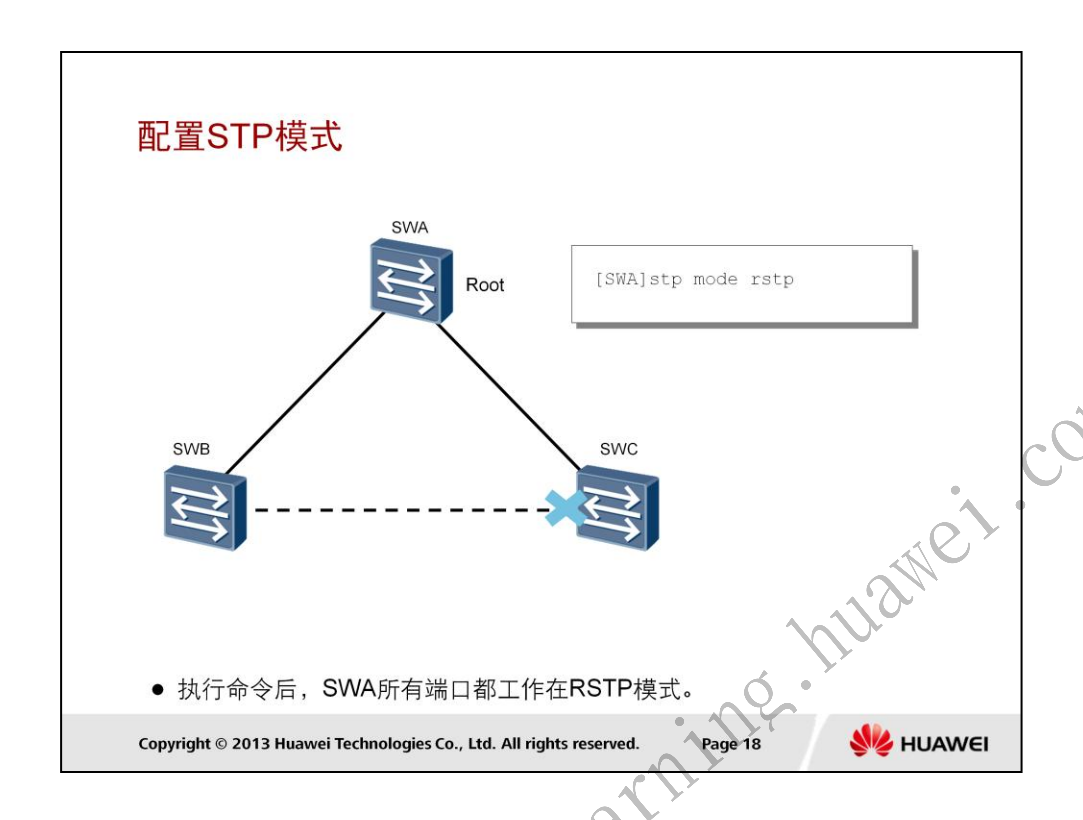
RSTP拓扑变化的处理类似于STP拓扑变化的处理，但也有些细微差别。

本例里面，SWC发生链路故障。SWA和SWC立即检测到链路故障并清除连接此链路的所有端口上的地址表项。在接下来的P/A进程中，交换机发送RST BPDU报文开始协商端口状态，拓扑变化通知报文也会随着Agreement置位的RST BPDU报文一起转发。RST BPDU报文里，Agreement和TC比特位都设置为1，通知上游交换机清除所有其他端口上的MAC地址表项，除了接收到TC置位的RST BPDU报文的端口。设置了TC位的RST BPDU报文周期性地转发给上游,在此周期时间内，所有相关接口上地址表项将会清除，接口上根据新的RSTP拓扑生成新的MAC地址表项。图形里面红色X表示由于拓扑变化导致端口上的MAC地址表项被清除。



RSTP是可以与STP实现后向兼容的，但在实际中，并不推荐这样的做法，原因是RSTP会失去其快速收敛的优势，而STP慢速收敛的缺点会暴露出来。

当同一个网段里既有运行STP的交换机又有运行RSTP的交换机时，STP交换机会忽略接收到的RST BPDU，而RSTP交换机在某端口上接收到STP BPDU时，会等待两个Hello Time时间之后，把自己的端口转换到STP工作模式，此后便发送STP BPDU，这样就实现了兼容性操作。



在Sx7交换机上，可以使用**stp mode rstp**命令来配置交换机工作在RSTP模式。

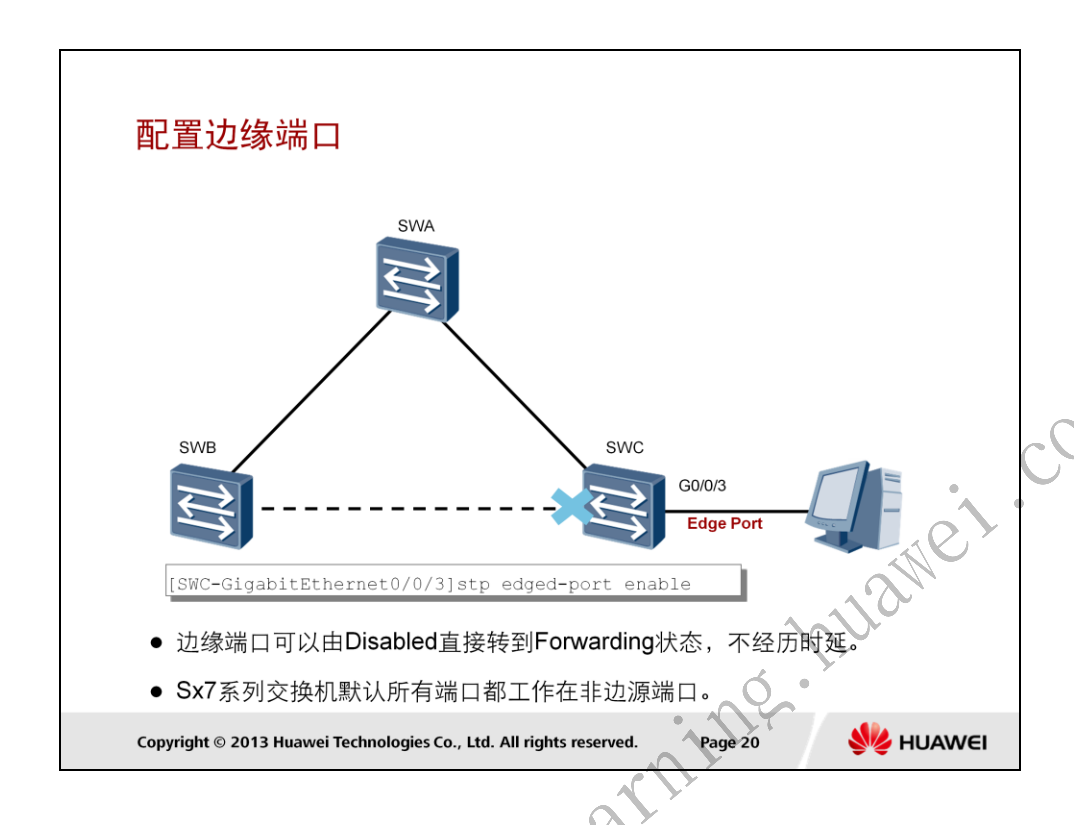
stp mode rstp命令在系统视图下执行，此命令必须在所有参与快速生成树拓扑计算的交换机上配置。

配置验证

```
[SWA]display stp
-------[CIST Global Info][Mode RSTP]-------
CIST Bridge         :32768 . 00-e0-fc-16-ee-43
Bridge Times        :Hello 2s MaxAge 20s FwDly 15s MaxHop 20
CIST Root/ERPC      :32768 . 00-e0-fc-16-ee-43 / 0
CIST RegRoot/IRPC   :32768 . 00-e0-fc-16-ee-43 / 0
CIST RootPortId     :0.0
BPDU-Protection     :Disabled
TC or TCN received  :37
TC count per hello  :0
STP Converge Mode   :Normal
Share region-configuration :Enabled
Time since last TC  :0 days 0h:14m:43s
```

 Page 19 HUAWEI

display stp命令可以显示RSTP配置信息和参数。根据显示信息可以确认交换机是否工作在RSTP模式。



边缘端口完全不参与STP或RSTP计算。边缘端口的状态要么是Disabled，要么是Forwarding；终端上电工作后，它就直接由Disabled状态转到Forwarding状态，终端下电后，它就直接由Forwarding状态转到Disabled状态。

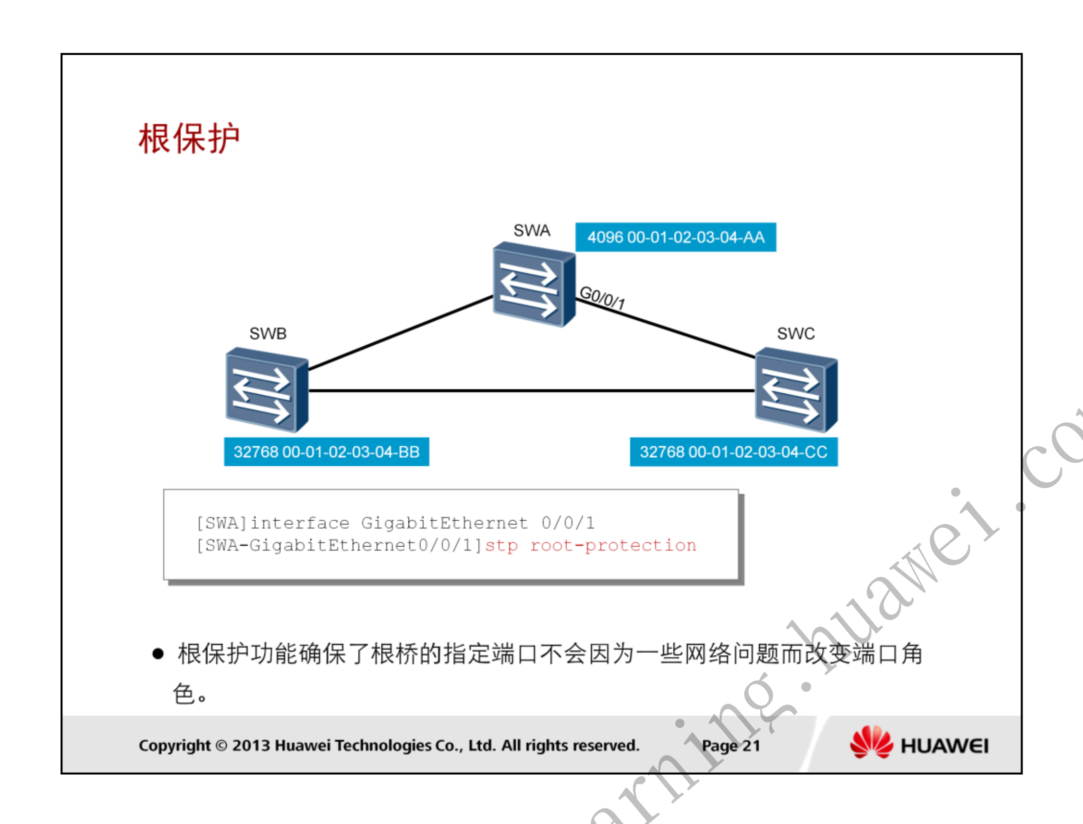
交换机所有端口默认为非边缘端口。

stp edged-port enable命令用来配置交换机的端口为边缘端口，它是一个针对某一具体端口的命令。

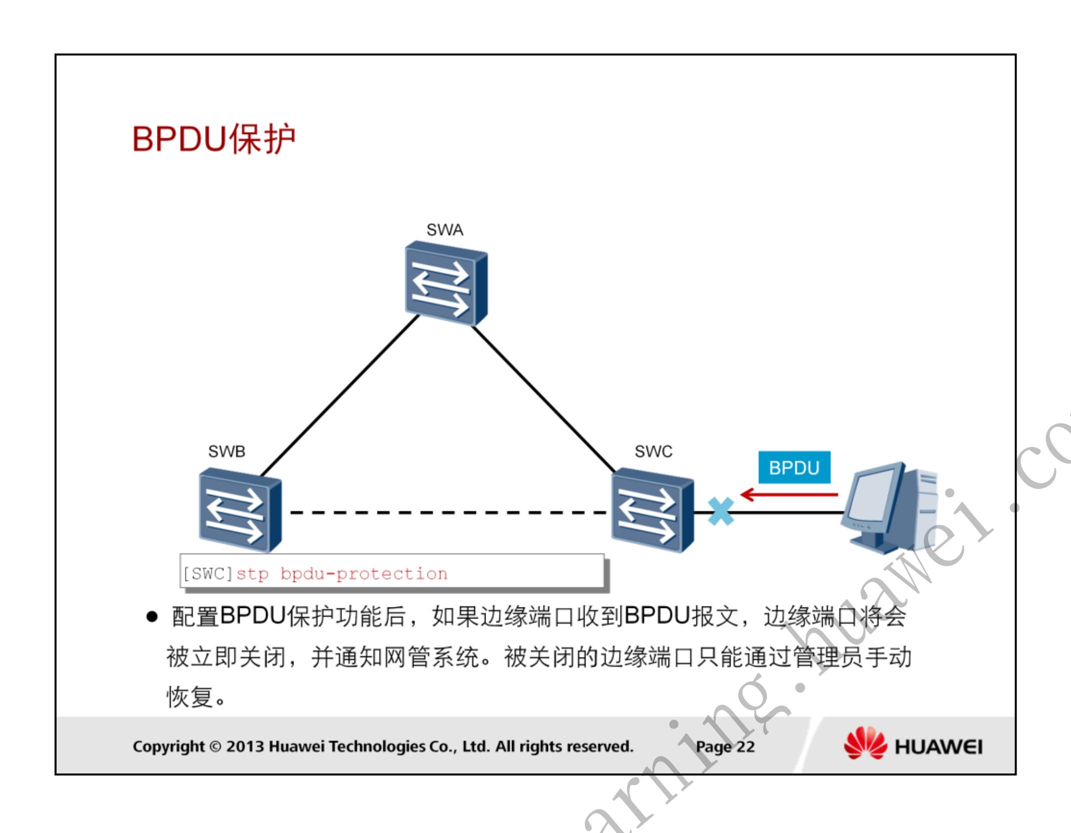
stp edged-port default命令用来配置交换机的所有端口为边缘端口。

stp edged-port disable命令用来将边缘端口的属性去掉，使之成为非边缘端口。它也是一个针对某一具体端口的命令。

需要注意的是，华为Sx7系列交换机运行STP时也可以使用边缘端口设置。



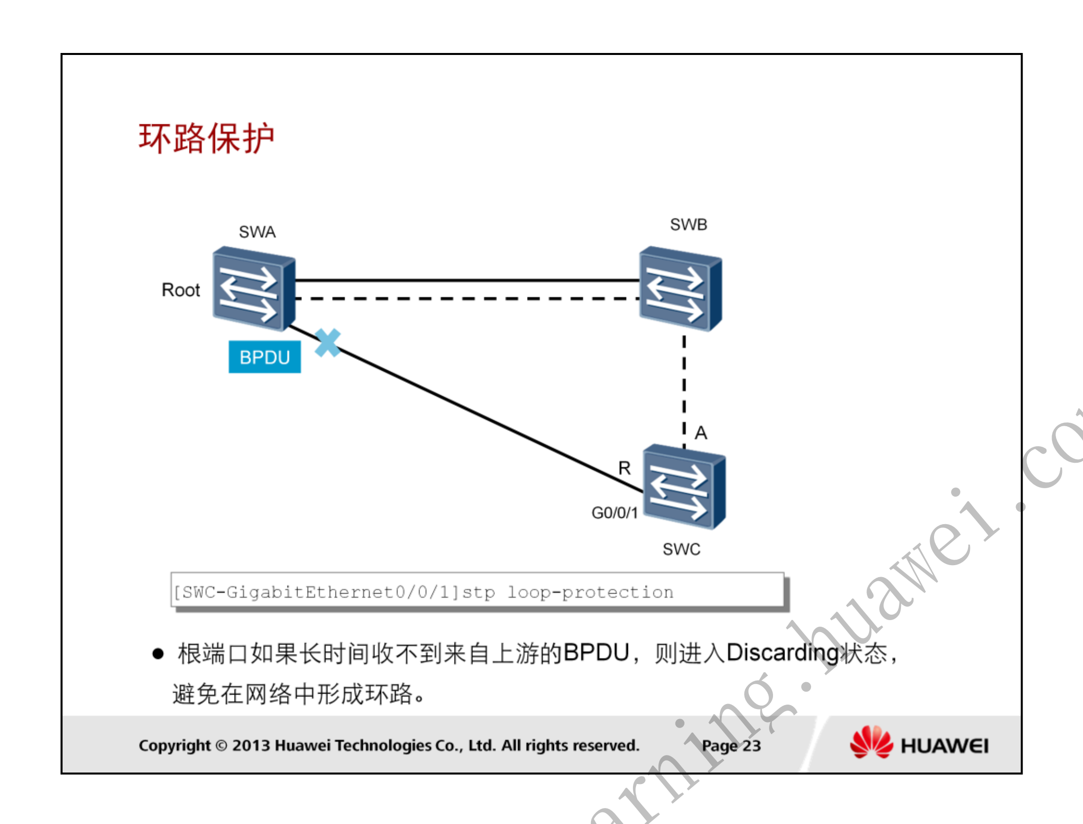
由于错误配置根交换机或网络中的恶意攻击，根交换机有可能会收到优先级更高的BPDU报文，使得根交换机变成非根交换机，从而引起网络拓扑结构的变动。这种不合法的拓扑变化，可能会导致原来应该通过高速链路的流量被牵引到低速链路上，造成网络拥塞。交换机提供了根保护功能来解决此问题。根保护功能通过维持指定端口角色从而保护根交换机。一旦启用了根保护功能的指定端口收到了优先级更高的BPDU报文时，端口会停止转发报文并且进入Listening状态。经过一段时间后，如果端口一直没有再收到优先级较高的BPDU报文，端口就会自动恢复到原来的状态。根保护功能仅在指定端口生效，不能配置在边缘端口或者使能了环路保护功能的端口上。



正常情况下，边缘端口是不会收到BPDU的。但是，如果有人发送BPDU来进行恶意攻击时，边缘端口就会收到这些BPDU，并自动变为非边缘端口，且开始参与网络拓扑计算，从而会增加整个网络的计算工作量，并可能引起网络震荡。

为防止上述情况的发生，我们可以使用BPDU保护功能。使能BPDU保护功能后的交换机的边缘端口在收到BPDU报文时，会立即关闭该端口，并通知网络管理系统。被关闭的边缘端口只能通过管理员手动进行恢复。

如需使能BPDU保护功能，可在系统视图下执行**stp bpdu-protection**命令。



交换机通过从上游交换机持续收到BPDU报文来维护根端口和阻塞端口的状态。当由于链路拥塞或者单向链路故障时，交换机不能收到上游交换机发送的BPDU报文，交换机重新选择根端口。最初的根端口会变成指定端口，阻塞端口进入Forwarding状态，这就有可能导致网络环路。

交换机提供了环路保护功能来避免这种环路的产生。环路保护功能使能后，如果根端口不能收到上游交换机发送的BPDU报文，则向网管发出通知信息。根端口会被阻塞，阻塞端口仍然将保持阻塞状态，这样就避免了可能发生的网络环路。

如需使能环路保护功能，可在接口视图下执行**stp loop-protection**命令。

配置验证

```
[SWC]display stp interface GigabitEthernet 0/0/1
----[CIST][Port1(GigabitEthernet0/0/1)][FORWARDING]----
 Port Protocol          :Enabled
 Port Role              :Designated Port
 Port Priority          :128
 Port Cost(Dot1T )      :Config=2000 / Active=2000
 Designated Bridge/Port    : 32768. 00-e0-fc-16-ee-43 / 128.1
 Port Edged             :Config=default / Active=disabled
 Point-to-point         :Config=auto / Active=true
 Transit Limit          :147 packets/hello-time
 Protection Type        :Loop
 Port STP Mode          :RSTP
 Port Protocol Type     :Config=auto / Active=dot1s
 BPDU Encapsulation     :Config=stp / Active=stp
……
```

 Page 24 HUAWEI

display stp interface *<interface>*命令可以显示端口的RSTP配置情况。包括端口状态，端口优先级，端口开销，端口角色，是否为边缘端口等等。



1. P/A机制是RSTP网络中的一种拓扑收敛机制，P/A进程中同步的作用是避免临时环路的产生。


谢谢
www.huawei.com

Module-4

网络间的互联


IP路由基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

以太网交换机工作在数据链路层，用于在网络内进行数据转发。而企业网络的拓扑结构一般会比较复杂，不同的部门，或者总部和分支可能处在不同的网络中，此时就需要使用路由器来连接不同的网络，实现网络之间的数据转发。



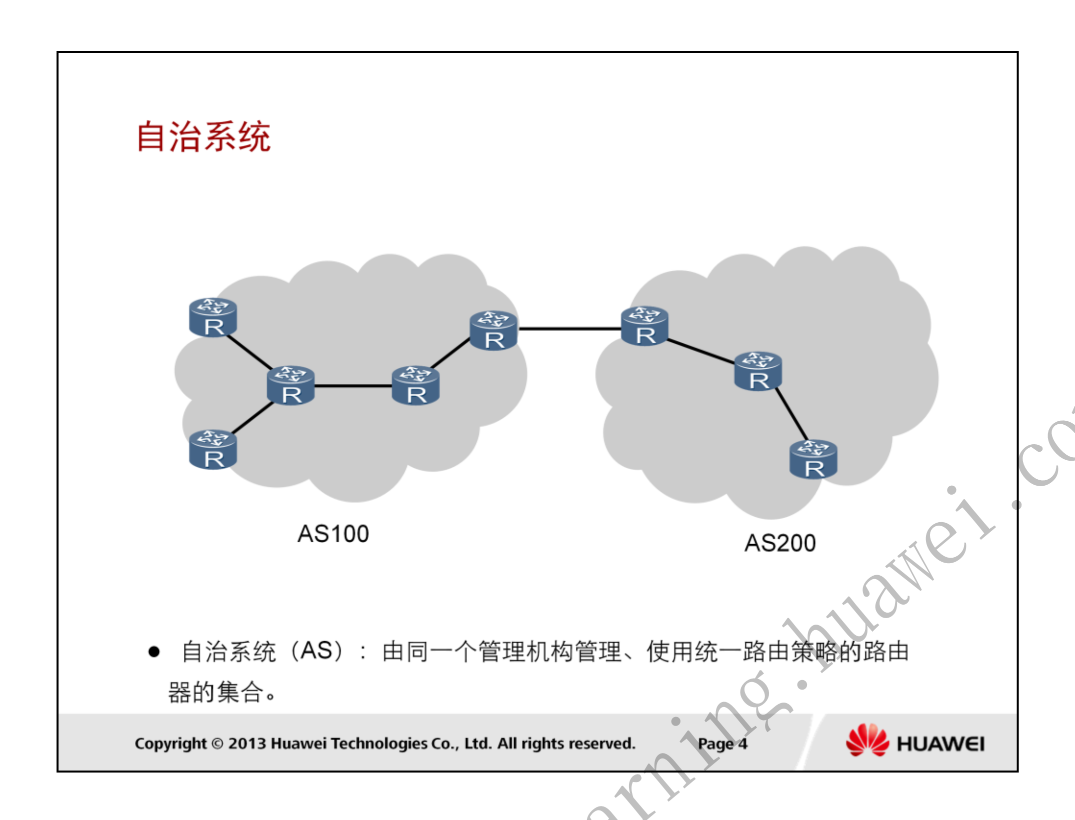


学习目标

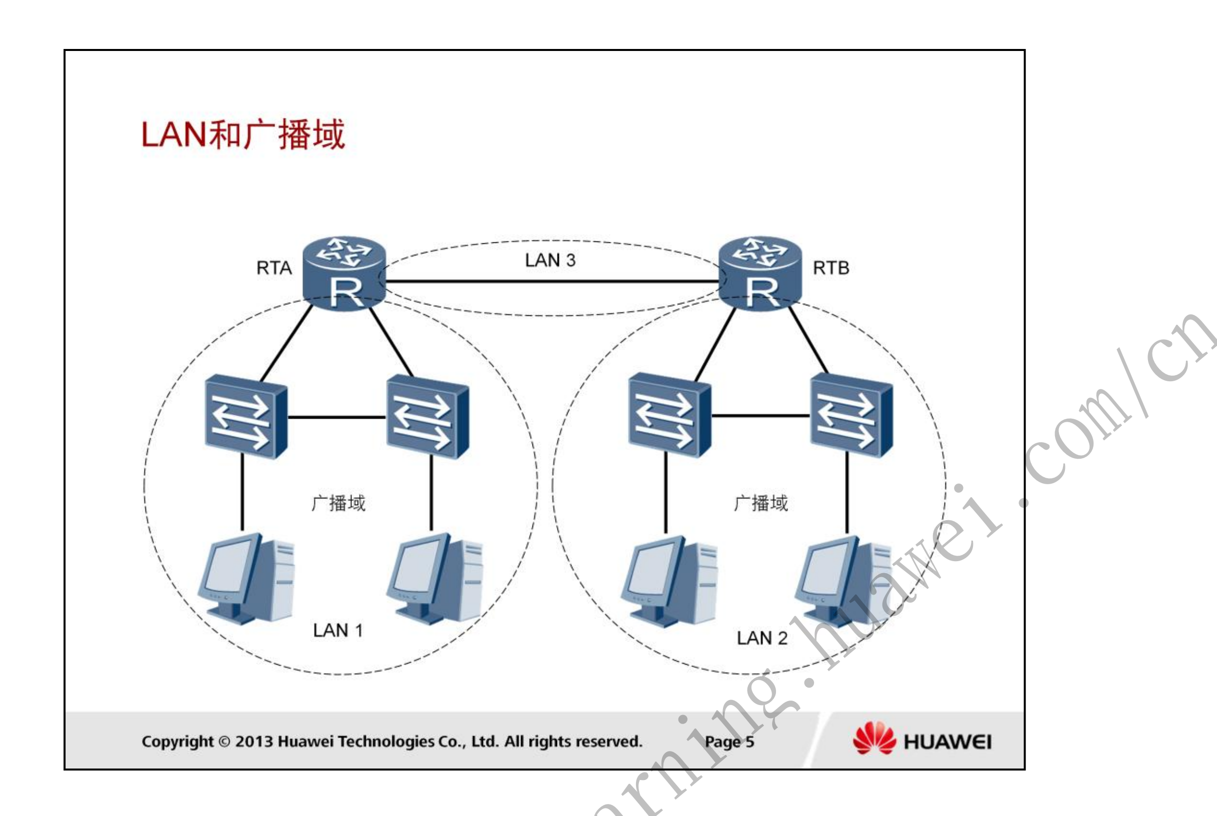
学完本课程后，您应该能：

- 掌握路由器的基本工作原理
- 掌握路由器选择最优路由的方法

  HUAWEI

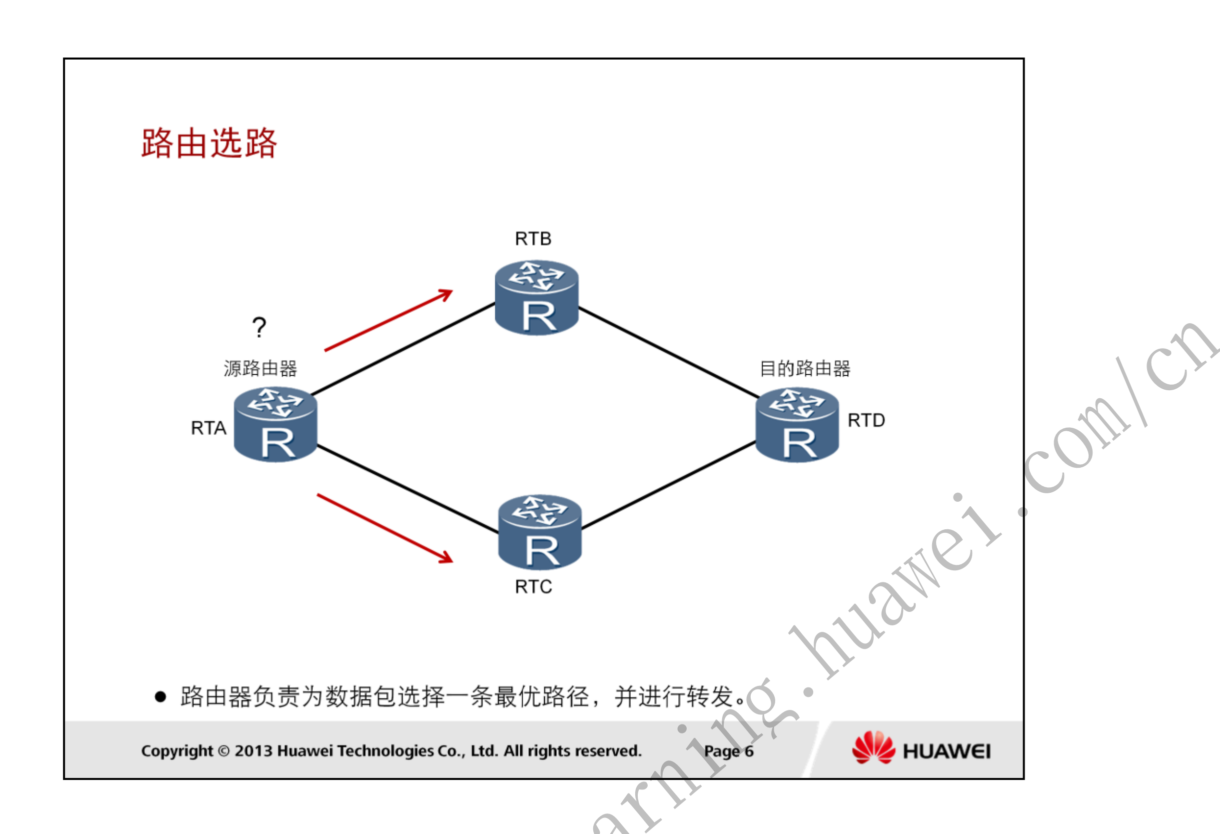


一般地我们可以把一个企业网络认为是一个自治系统AS（Autonomous System）。根据RFC1030的定义，自治系统是由一个单一实体管辖的网络，这个实体可以是一个互联网服务提供商，或一个大型组织机构。自治系统内部遵循一个单一且明确的路由策略。最初，自治系统内部只考虑运行单个路由协议；然而，随着网络的发展，一个自治系统内现在也可以支持同时运行多种路由协议。



一个AS通常由多个不同的局域网组成。以企业网络为例，各个部门可以属于不同的局域网，或者各个分支机构和总部也可以属于不同的局域网。局域网内的主机可以通过交换机来实现相互通信。不同局域网之间的主机要想相互通信，可以通过路由器来实现。路由器工作在网络层，隔离了广播域，并可以作为每个局域网的网关，发现到达目的网络的最优路径，最终实现报文在不同网络间的转发。

此例中，RTA和RTB把整个网络分成了三个不同的局域网，每个局域网为一个广播域。LAN1内部的主机直接可以通过交换机实现相互通信，LAN2内部的主机之间也是如此。但是，LAN1内部的主机与LAN2内部的主机之间则必须要通过路由器才能实现相互通信。



路由器收到数据包后，会根据数据包中的目的IP地址选择一条最优的路径，并将数据包转发到下一个路由器，路径上最后的路由器负责将数据包送交目的主机。数据包在网络上的传输就好像是体育运动中的接力赛一样，每一个路由器负责将数据包按照最优的路径向下一跳路由器进行转发，通过多个路由器一站一站的接力，最终将数据包通过最优路径转发到目的地。当然有时候由于实施了一些特别的路由策略，数据包通过的路径可能并不一定是最佳的。

路由器能够决定数据报文的转发路径。如果有多条路径可以到达目的地，则路由器会通过进行计算来决定最佳下一跳。计算的原则会随实际使用的路由协议不同而不同。

IP路由表

```
[Huawei]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public  Destinations : 2        Routes : 2
Destination/Mask  Proto  Pre  Cost  Flags  NextHop   Interface
0.0.0.0/0         Static 60   0     D      120.0.0.2 Serial1/0/0
8.0.0.0/8         RIP    100  3     D      120.0.0.2 Serial1/0/0
9.0.0.0/8         OSPF   10   50    D      20.0.0.2  Ethernet2/0/0
9.1.0.0/16        RIP    100  4     D      120.0.0.2 Serial1/0/0
11.0.0.0/8        Static 60   0     D      120.0.0.2 Serial2/0/0
20.0.0.0/8        Direct 0    0     D      20.0.0.1  Ethernet2/0/0
20.0.0.1/32       Direct 0    0     D      127.0.0.1 LoopBack0
```

- 路由表中包含了路由器可以到达的目的网络。目的网络在路由表中不存在的数据包会被丢弃。

 Page 7



路由器转发数据包的关键是路由表。每个路由器中都保存着一张路由表，表中每条路由项都指明了数据包要到达某网络或某主机应通过路由器的哪个物理接口发送，以及可到达该路径的哪个下一个路由器，或者不再经过别的路由器而直接可以到达目的地。

路由表中包含了下列关键项：

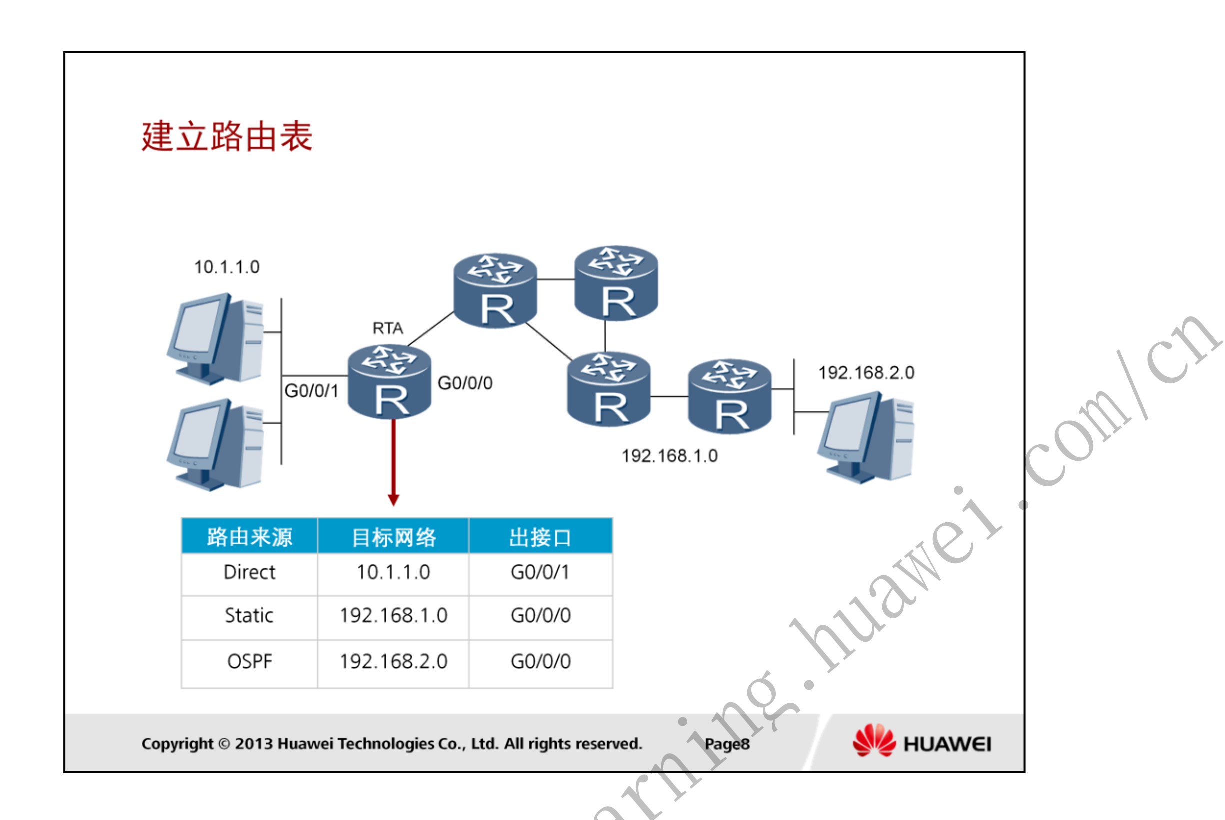
目的地址（Destination）：用来标识IP包的目的地址或目的网络。

网络掩码（Mask）：在IP编址课程中已经介绍了网络掩码的结构和作用。同样，在路由表中网络掩码也具有重要的意义。IP地址和网络掩码进行“逻辑与”便可得到相应的网段信息。如本例中：目的地址为8.0.0.0，掩码为255.0.0.0，相与后便可得到一个A类的网段信息(8.0.0.0/8)。网络掩码的另一个作用还表现在当路由表中有多条目的地址相同的路由信息时，路由器将选择其掩码最长的一项作为匹配项。

输出接口（Interface）：指明IP包将从该路由器的哪个接口转发出去。

下一跳IP地址（NextHop）：指明IP包所经由的下一个路由器的接口地址。

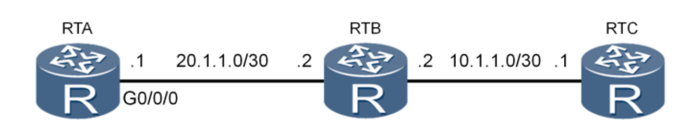
路由表中优先级、度量值等其他的几个字段我们将在以后进行介绍。



根据来源的不同，路由表中的路由通常可分为以下三类：

1. 链路层协议发现的路由（也称为接口路由或直连路由）。
2. 由网络管理员手工配置的静态路由。
3. 动态路由协议发现的路由。

最长匹配原则



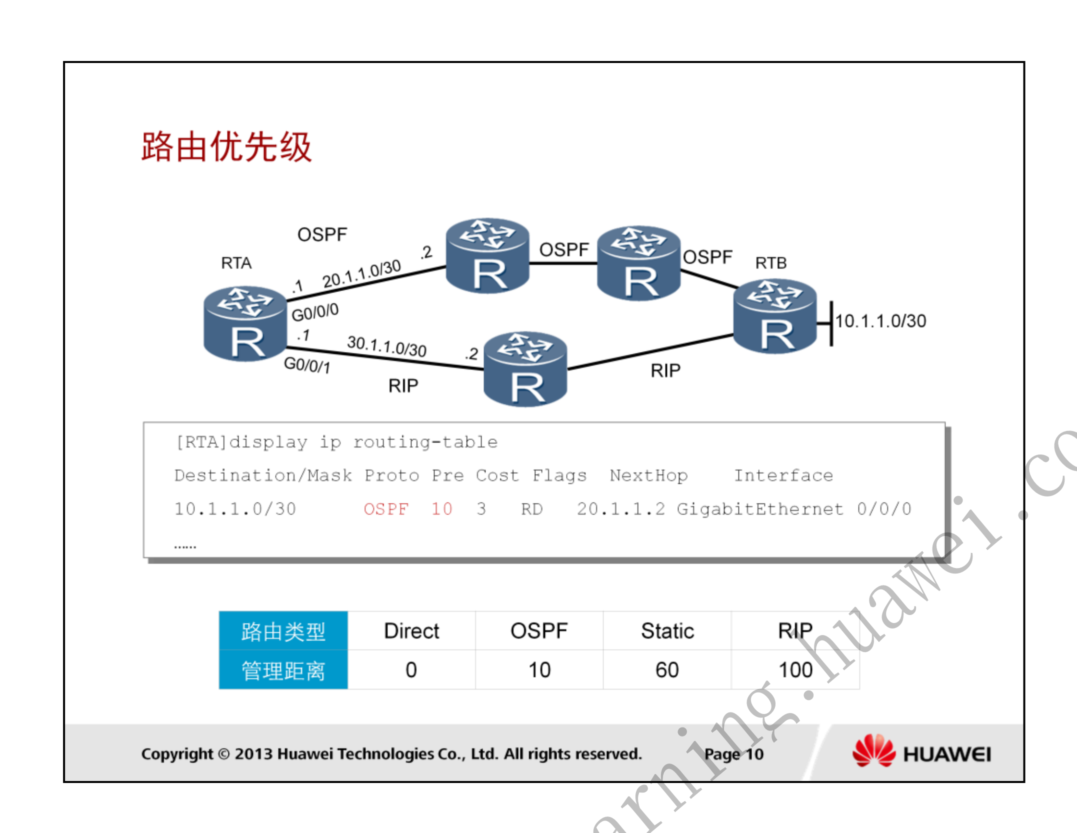

```
[RTA]display ip routing-table
Destination/Mask Proto  Pre  Cost Flags NextHop    Interface
10.1.1.0/24      Static  60   0    RD   20.1.1.2 GigabitEthernet 0/0/0
10.1.1.0/30      Static  60   0    RD   20.1.1.2 GigabitEthernet 0/0/0
```

- 路由表中如果有多个匹配目的网络的路由条目，则路由器会选择掩码最长的条目。

 Page 9 

路由器在转发数据时，需要选择路由表中的最优路由。当数据报文到达路由器时，路由器首先提取出报文的目的IP地址，然后查找路由表，将报文的目的IP地址与路由表中某表项的掩码字段做“与”操作，“与”操作后的结果跟路由表该表项的目的IP地址比较，相同则匹配上，否则就没有匹配上。 当与所有的路由表项都进行匹配后，路由器会选择一个掩码最长的匹配项。

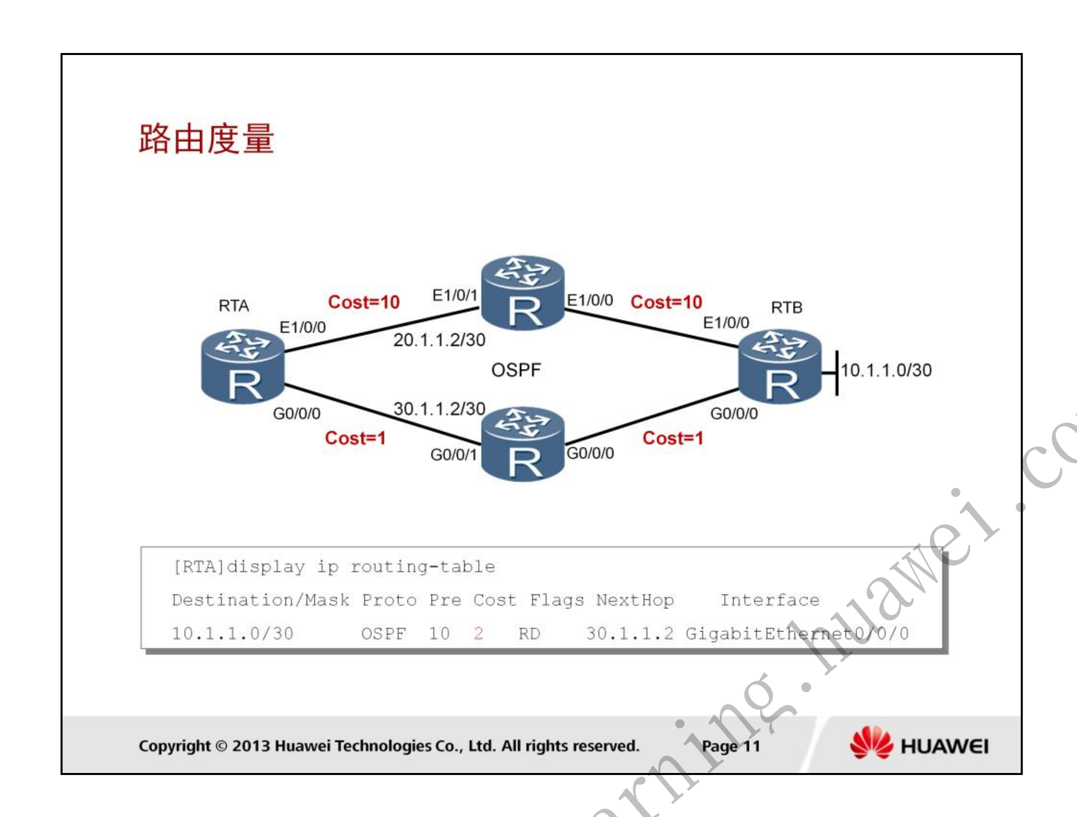
如图所示，路由表中有两个表项到达目的网段10.1.1.0，下一跳地址都是20.1.1.2。如果要将报文转发至网段10.1.1.1，则10.1.1.0/30符合最长匹配原则。



路由器可以通过多种不同协议学习到去往同一目的网络的路由，当这些路由都符合最长匹配原则时，必须决定哪个路由优先。

每个路由协议都有一个协议优先级（取值越小、优先级越高）。当有多个路由信息时，选择最高优先级的路由作为最佳路由。

如图所示，路由器通过两种路由协议学习到了网段10.1.1.0的路由。虽然RIP协议提供了一条看起来更加直连的路线，但是由于OSPF具有更高的优先级，因而成为优选路由，并被加入路由表中。



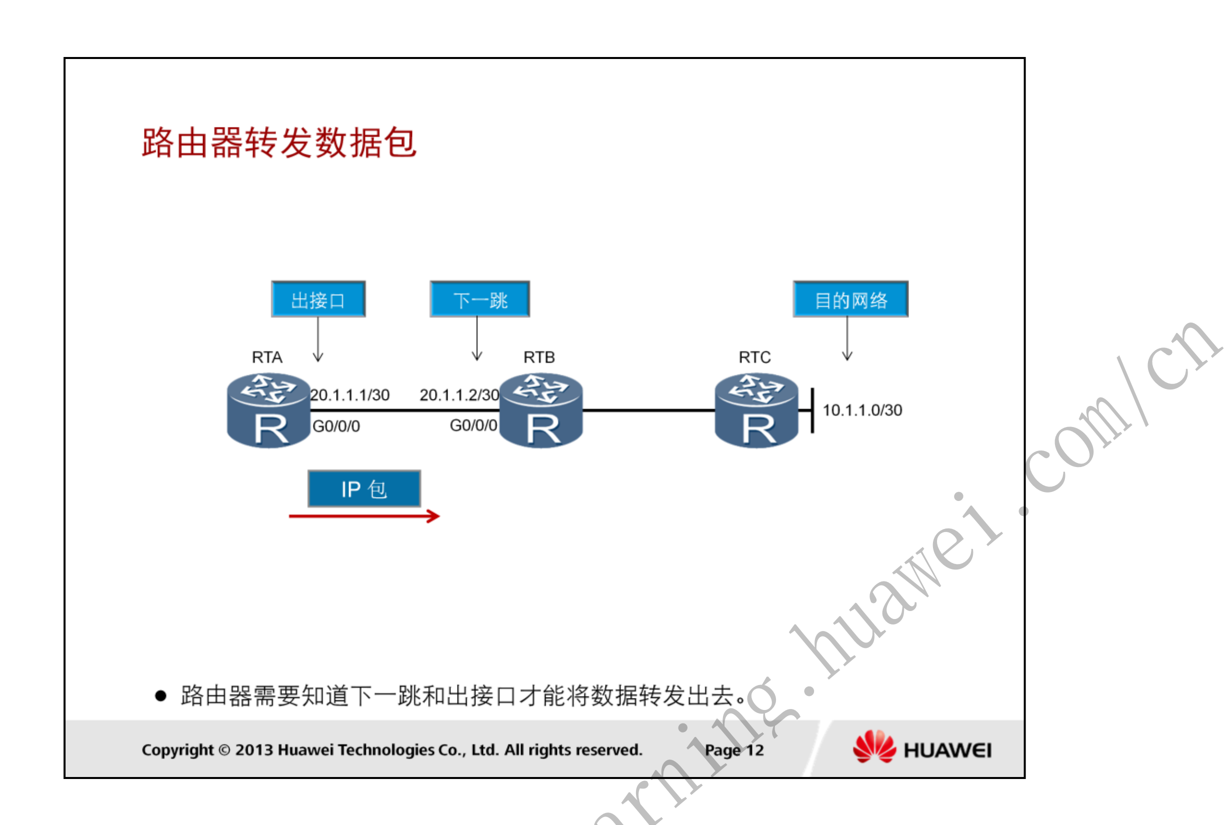
如果路由器无法用优先级来判断最优路由，则使用度量值（metric）来决定需要加入路由表的路由。

一些常用的度量值有：跳数，带宽，时延，代价，负载，可靠性等。

跳数是指到达目的地所通过的路由器数目。

带宽是指链路的容量，高速链路开销（度量值）较小。

metric值越小，路由越优先；因此，图示中metric=1+1=2的路由是到达目的地的最优路由，其表项可以在路由表中找到。



路由器收到一个数据包后，会检查其目的IP地址，然后查找路由表。查找到匹配的路由表项之后，路由器会根据该表项所指示的出接口信息和下一跳信息将数据包转发出去。



总结

- 路由器选择最优路由的顺序是什么?
- Preference字段在路由表中代表什么含义?



1. 路由器在选择最优路由时，会首先比较相同路由的优先级；如果优先级相同，再比较metric值。如果metric值也相同，多条相同的路由将实现等价负载分担。
2. Preference字段在路由表中代表了路由优先级。设备厂商会在各自的产品中为不同的路由协议规定不同的优先级。


谢谢
www.huawei.com


静态路由基础
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

静态路由是指由管理员手动配置和维护的路由。

静态路由配置简单，被广泛应用于网络中。另外，静态路由还可以实现负载均衡和路由备份。因此，学习并掌握好静态路由的应用与配置是非常必要的。

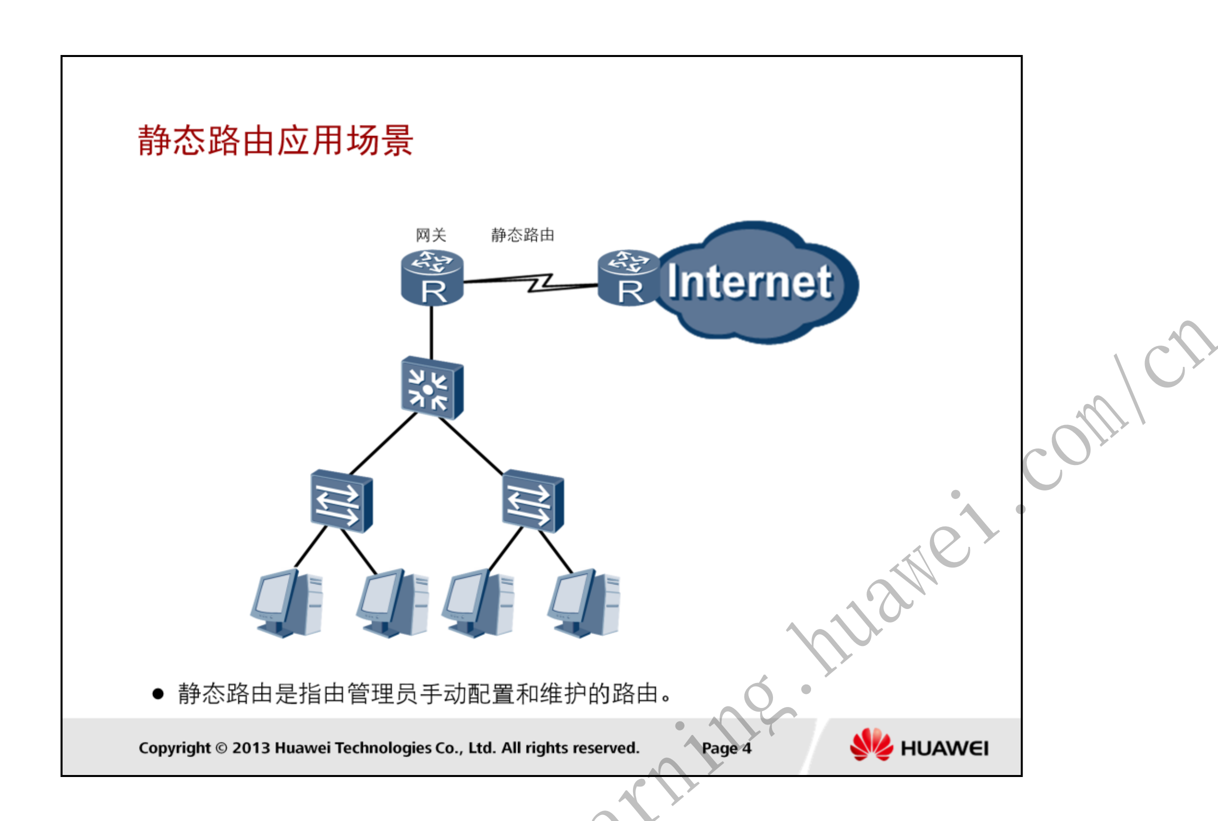




学习目标

学完本课程后，您应该能：

- 识别静态路由的应用场景
- 掌握静态路由的配置



静态路由是指由管理员手动配置和维护的路由。静态路由配置简单，并且无需像动态路由那样占用路由器的CPU资源来计算和分析路由更新。

静态路由的缺点在于，当网络拓扑发生变化时，静态路由不会自动适应拓扑改变，而是需要管理员手动进行调整。

静态路由一般适用于结构简单的网络。在复杂网络环境中，一般会使用动态路由协议来生成动态路由。不过，即使是在复杂网络环境中，合理地配置一些静态路由也可以改进网络的性能。

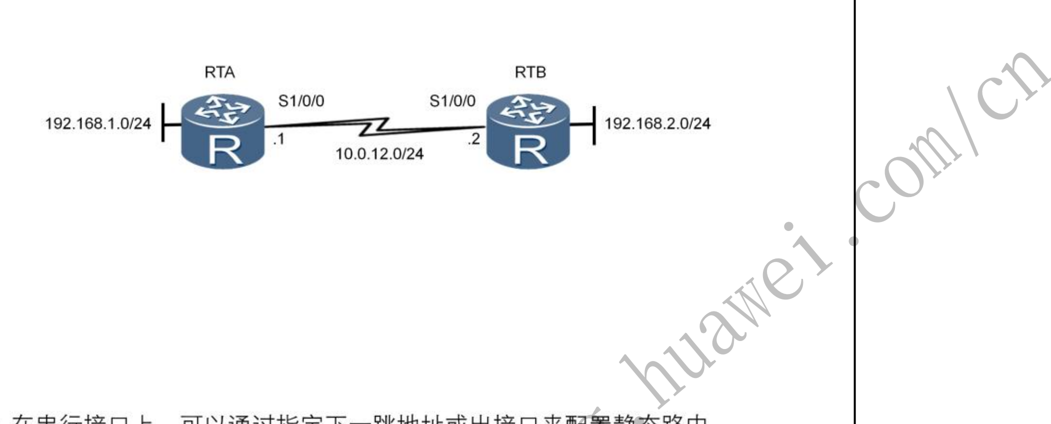
静态路由配置


```
[RTB]ip route-static 192.168.1.0 255.255.255.0 10.0.12.1
[RTB]ip route-static 192.168.1.0 255.255.255.0 Serial 1/0/0
[RTB]ip route-static 192.168.1.0 24 Serial 1/0/0
```

 Page 5 HUAWEI

ip route-static *ip-address* { *mask* | *mask-length* } *interface-type interface-number* [*nexthop-address*]命令用来配置静态路由。参数*ip-address*指定了一个网络或者主机的目的地址，参数*mask*指定了一个子网掩码或者前缀长度。如果使用了广播接口如以太网接口作为出接口，则必须要指定下一跳地址；如果使用了串口作为出接口，则可以通过参数*interface-type*和*interface-number*（如Serial 1/0/0）来配置出接口，此时不必指定下一跳地址。

静态路由



- 在串行接口上，可以通过指定下一跳地址或出接口来配置静态路由。





静态路由可以应用在串行网络或以太网中，但静态路由在这两种网络中的配置有所不同。

在串行网络中配置静态路由时，可以只指定下一跳地址或只指定出接口。华为ARG3系列路由器中，串行接口默认封装PPP协议，对于这种类型的接口，静态路由的下一跳地址就是与接口相连的对端接口的地址，所以在串行网络中配置静态路由时可以只配置出接口。

以太网是广播类型网络，和串行网络情况不同。在以太网中配置静态路由，必须指定下一跳地址。

静态路由

RTA
192.168.1.0/24
.1
10.0.123.0/24
.2
.3
192.168.2.0/24
RTB
192.168.3.0/24
RTC

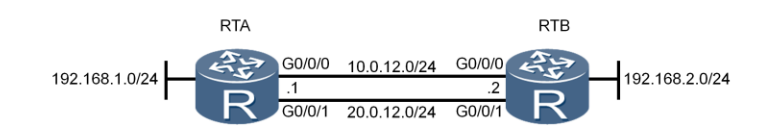
- 在广播型的接口（如以太网接口）上配置静态路由时，必须要指定下一跳地址。

 Page 7



在广播型的接口上配置静态路由时，必须明确指定下一跳地址。以太网中同一网络可能连接了多台路由器，如果在配置静态路由时只指定了出接口，则路由器无法将报文转发到正确的下一跳。在本示例中，RTA需要将数据转发到192.168.2.0/24网络，在配置静态路由时，需要明确指定下一跳地址为10.0.123.2，否则，RTA将无法将报文转发到RTB所连接的192.168.2.0/24网络，因为RTA不知道应该通过RTB还是RTC才能到达目的地。

负载分担




```
[RTB]ip route-static 192.168.1.0 255.255.255.0 10.0.12.1
[RTB]ip route-static 192.168.1.0 255.255.255.0 20.0.12.1
```

- 静态路由支持到达同一目的地的等价负载分担。

 Page 8



当源网络和目的网络之间存在多条链路时，可以通过等价路由来实现流量负载分担。这些等价路由具有相同的目的网络和掩码、优先级和度量值。

本示例中RTA和RTB之间有两条链路相连，通过使用等价的静态路由来实现流量负载分担。

在RTB上配置了两条静态路由，它们具有相同的目的IP地址和子网掩码、优先级（都为60）、路由开销（都为0），但下一跳不同。在RTB需要转发数据给RTA时，就会使用这两条等价静态路由将数据进行负载分担。

在RTA上也应该配置对应的两条等价的静态路由。

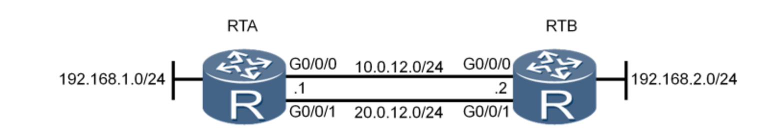
配置验证

```
[RTB]display ip routing-table
Route Flags: R - relay, D - download to fib
--------------------------------------------------------------
Routing Tables: Public  Destinations : 13       Routes : 14
Destination/Mask  Proto Pre Cost Flags NextHop Interface
……
192.168.1.0/24     Static 60  0   RD 10.0.12.1 GigabitEthernet 0/0/0
                   Static 60  0   RD 20.0.12.1 GigabitEthernet 0/0/1
```

 Page 9 HUAWEI

在配置完静态路由之后，可以使用**display ip routing-table**命令来验证配置结果。在本示例中，红色高亮部分代表路由表中的静态路由。这两条路由具有相同的目的地址和掩码，并且有相同的优先级和度量值，但是它们的下一跳地址和出接口不同。此时，RTB就可以通过这两条等价路由实现负载分担。

路由备份




```
[RTB]ip route-static 192.168.1.0 255.255.255.0 10.0.12.1
[RTB]ip route-static 192.168.1.0 255.255.255.0 20.0.12.1 preference 100
```

- 浮动静态路由在网络中主路由失效的情况下，会加入到路由表并承担数据转发业务。

 Page 10



在配置多条静态路由时，可以修改静态路由的优先级，使一条静态路由的优先级高于其他静态路由，从而实现静态路由的备份，也叫浮动静态路由。在本示例中，RTB上配置了两条静态路由。正常情况下，这两条静态路由是等价的。通过配置preference 100，使第二条静态路由的优先级要低于第一条（值越大优先级越低）。路由器只把优先级最高的静态路由加入到路由表中。当加入到路由表中静态路由出现故障时，优先级低的静态路由才会加入到路由表并承担数据转发业务。

配置验证

```
[RTB]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------
Routing Tables: Public  Destinations : 13       Routes : 14
Destination/Mask Proto Pre Cost Flags NextHop Interface
……
192.168.1.0/24  Static  60   0  RD  10.0.12.1 GigabitEthernet0/0/0
```

- 在主链路正常情况下，只有主路由会出现在路由表中。



从**display ip routing-table**命令的回显信息中可以看出，通过修改静态路由优先级实现了浮动静态路由。正常情况下，路由表中应该显示两条有相同目的地、但不同下一跳和出接口的等价路由。由于修改了优先级，回显中只有一条默认优先级为60的静态路由。另一条静态路由的优先级是100，该路由优先级低，所以不会显示在路由表中。

配置验证

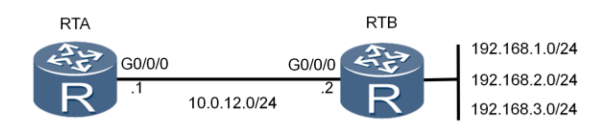
```
[RTB]interface GigabitEthernet 0/0/0
[RTB-GigabitEthernet 0/0/0]shutdown
[RTB]display ip routing-table
Route Flags: R - relay, D - download to fib
--------------------------------------------------------------
Routing Tables: Public  Destinations : 13       Routes : 14
Destination/Mask Proto Pre Cost Flags NextHop Interface
……
192.168.1.0/24 Static  100  0  RD  20.0.12.1 GigabitEthernet 0/0/1
```

- 在主链路出现故障时，浮动静态路由会被激活并加入到路由表中，承担数据转发业务。

 Page 12 

当主用静态路由出现物理链路故障或者接口故障时，该静态路由不能再提供到达目的地的路径，所以在路由表中会被删除。此时，浮动静态路由会被加入到路由表，以保证报文能够从备份链路成功转发到目的地。在主用静态路由的物理链路恢复正常后，主用静态路由会重新被加入到路由表，并且数据转发业务会从浮动静态路由切换到主用静态路由，而浮动静态路由会在路由表中再次被隐藏。

缺省路由




```
[RTA]ip route-static 0.0.0.0 0.0.0.0 10.0.12.2
[RTA]ip route-static 0.0.0.0 0 10.0.12.2 GigabitEthernet 0/0/0
```

- 缺省路由是目的地址和掩码都为全0的特殊路由。
- 如果报文的目的地址无法匹配路由表中的任何一项，路由器将选择依照缺省路由来转发报文。

 Page 13 

当路由表中没有与报文的目的地址匹配的表项时，设备可以选择缺省路由作为报文的转发路径。在路由表中，缺省路由的目的网络地址为0.0.0.0，掩码也为0.0.0.0。在本示例中，RTA使用缺省路由转发到达未知目的地址的报文。缺省静态路由的默认优先级也是60。在路由选择过程中，缺省路由会被最后匹配。

配置验证

```
[RTA]display ip routing-table
Route Flags: R - relay, D - download to fib
--------------------------------------------------------------
Routing Tables: Public  Destinations : 13      Routes : 14
Destination/Mask Proto Pre Cost Flags NextHop Interface
……
 0.0.0.0/0      Static  60   0  RD   10.0.12.2 GigabitEthernet0/0/0
```

 Page 14 HUAWEI

配置缺省路由后，可以使用**display ip routing-table**命令来查看该路由的详细信息。在本示例中，目的地址在路由表中没能匹配的所有报文都将通过GigabitEthernet 0/0/0接口转发到下一跳地址10.0.12.2。



总结

- 如何配置能够将静态路由配置为浮动静态路由?
- 配置缺省路由时，目的网络地址是什么?

 Page 15



1. 在配置静态路由时，需要调整其中一条静态路由的优先级，就可将其修改为浮动静态路由。
2. 在配置缺省路由时，目的网络为0.0.0.0，代表的是任意网络。


谢谢
www.huawei.com


距离矢量路由协议-RIP
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

路由信息协议RIP（Routing Information Protocol）的简称，它是一种基于距离矢量（Distance-Vector）算法的协议，使用跳数作为度量来衡量到达目的网络的距离。RIP主要应用于规模较小的网络中。

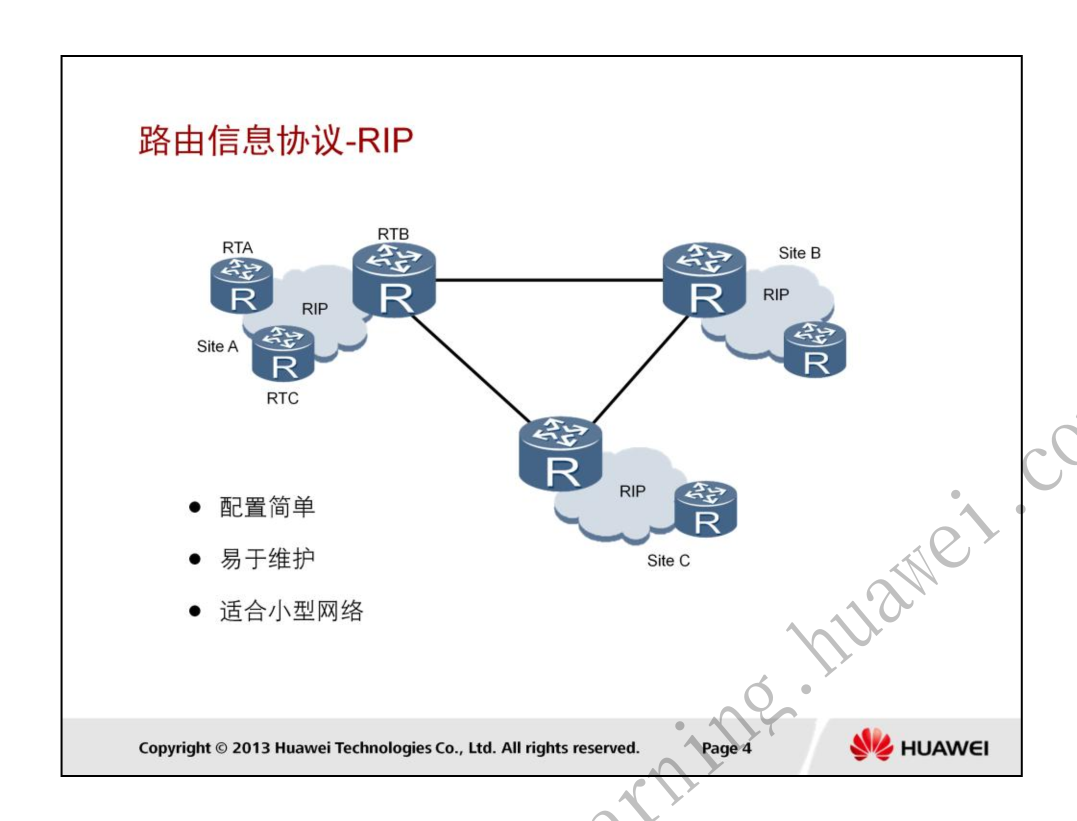




学习目标

学完本课程后，您应该能：

- 掌握RIP的基本工作原理
- 掌握RIP的配置

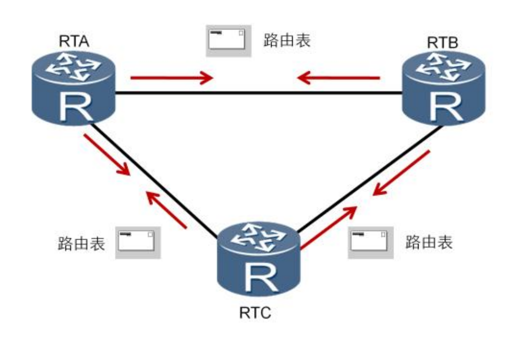


RIP是一种比较简单的内部网关协议。RIP使用了基于距离矢量的贝尔曼-福特算法（Bellman-Ford）来计算到达目的网络的最佳路径。

最初的RIP协议开发时间较早，所以在带宽、配置和管理方面要求也较低，因此，RIP主要适合于规模较小的网络中。

RIP协议中定义的相关参数也比较少。例如，它不支持VLSM和CIDR，也不支持认证功能。

RIP工作原理



- 路由器运行RIP后，会首先发送路由更新请求，收到请求的路由器会发送自己的RIP路由进行响应。
- 网络稳定后，路由器会周期性发送路由更新信息。

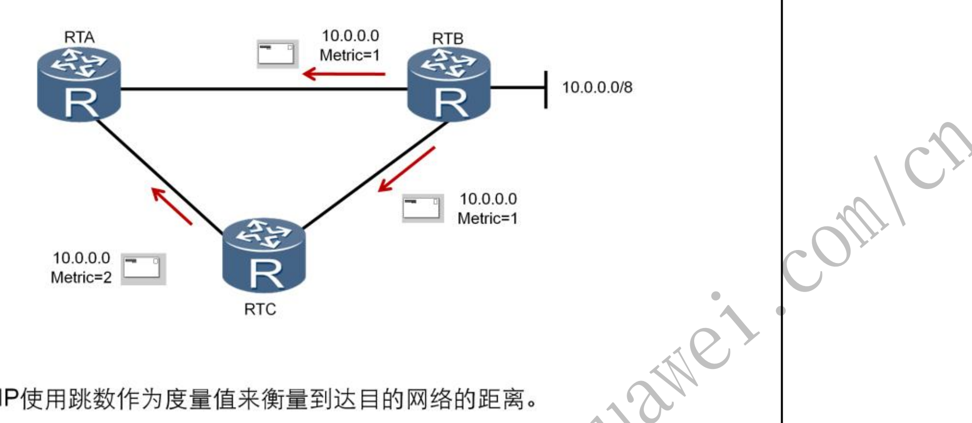
 Page 5



路由器启动时，路由表中只会包含直连路由。运行RIP之后，路由器会发送Request报文，用来请求邻居路由器的RIP路由。运行RIP的邻居路由器收到该Request报文后，会根据自己的路由表，生成Response报文进行回复。路由器在收到Response报文后，会将相应的路由添加到自己的路由表中。

RIP网络稳定以后，每个路由器会周期性地向邻居路由器通告自己的整张路由表中的路由信息，默认周期为30秒。邻居路由器根据收到的路由信息刷新自己的路由表。

RIP-度量



- RIP使用跳数作为度量值来衡量到达目的网络的距离。
- 缺省情况下，直连网络的路由跳数为0。当路由器发送路由更新时，会把度量值加1。RIP规定超过15跳为网络不可达。

 Page 6



RIP使用跳数作为度量值来衡量到达目的网络的距离。在RIP中，路由器到与它直接相连网络的跳数为0，每经过一个路由器后跳数加1。为限制收敛时间，RIP规定跳数的取值范围为0~15之间的整数，大于15的跳数被定义为无穷大，即目的网络或主机不可达。

路由器从某一邻居路由器收到路由更新报文时，将根据以下原则更新本路由器的RIP 路由表：

1. 对于本路由表中已有的路由项，当该路由项的下一跳是该邻居路由器时，不论度量值将增大或是减少，都更新该路由项（度量值相同时只将其老化定时器清零。路由表中的每一路由项都对应了一个老化定时器，当路由项在 180 秒内没有任何更新时，定时器超时，该路由项的度量值变为不可达）。
2. 当该路由项的下一跳不是该邻居路由器时，如果度量值将减少，则更新该路由项。
3. 对于本路由表中不存在的路由项，如果度量值小于16，则在路由表中增加该路由项。

某路由项的度量值变为不可达后，该路由会在 Response 报文中发布四次（120 秒），然后从路由表中清除。

在本示例中，路由器RTA通过两个接口学习路由信息，每条路由信息都有相应的度量值，到达目的网络的最佳路由就是通过这些度量值计算出来的。

RIPv1 vs. RIPv2

- RIPv1是有类别路由协议，不支持VLSM和CIDR。
- 以广播的形式发送报文。
- 不支持认证。

- RIPv2为无类别路由协议，支持VLSM，支持路由聚合与CIDR。
- 支持以广播或者组播（224.0.0.9）方式发送报文。
- 支持明文认证和 MD5 密文认证。

 Page7 

RIP包括RIPv1和RIPv2两个版本。

RIPv1为有类别路由协议，不支持VLSM和CIDR。 RIPv2为无类别路由协议，支持VLSM，支持路由聚合与CIDR。

RIPv1使用广播发送报文；RIPv2有两种发送方式：广播方式和组播方式，缺省是组播方式。RIPv2的组播地址为224.0.0.9。组播发送报文的好处是在同一网络中那些没有运行RIP的网段可以避免接收RIP的广播报文；另外，组播发送报文还可以使运行RIPv1的网段避免错误地接收和处理RIPv2中带有子网掩码的路由。

RIPv1不支持认证功能，RIPv2支持明文认证和MD5密文认证。

RIPv1报文格式

UDP 端口：520

目的地址：255.255.255.255

RTA G0/0/0 —— G0/0/0 RTB

Command	Version	Must be Zero
Address Family Identifier		Must be Zero
IP Address		
Must be Zero		
Must be Zero		
Metric		

 Page 8 HUAWEI

RIP协议通过UDP交换路由信息，端口号为520。RIPv1以广播形式发送路由信息，目的IP地址为广播地址255.255.255.255。

报文格式中每个字段的值和作用：

1. Command：表示该报文是一个请求报文还是响应报文，只能取1或者2。1表示该报文是请求报文，2表示该报文是响应报文。
2. Version：表示RIP的版本信息。对于RIPv1，该字段的值为1。
3. Address Family Identifier（AFI）：表示地址标识信息，对于IP协议，其值为2。
4. IP address：表示该路由条目的目的IP地址。这一项可以是网络地址、主机地址。
5. Metric：标识该路由条目的度量值，取值范围1-16。

一个RIP路由更新消息中最多可包含25条路由表项，每个路由表项都携带了目的网络的地址和度量值。整个RIP报文大小限制为不超过504字节。如果整个路由表的更新消息超过该大小，需要发送多个RIPv1报文。

RIPv2报文格式

Command	Version	Unused
Address Family Identifier		Route Tag
IP Address		
Subnet Mask		
Next Hop		
Metric		

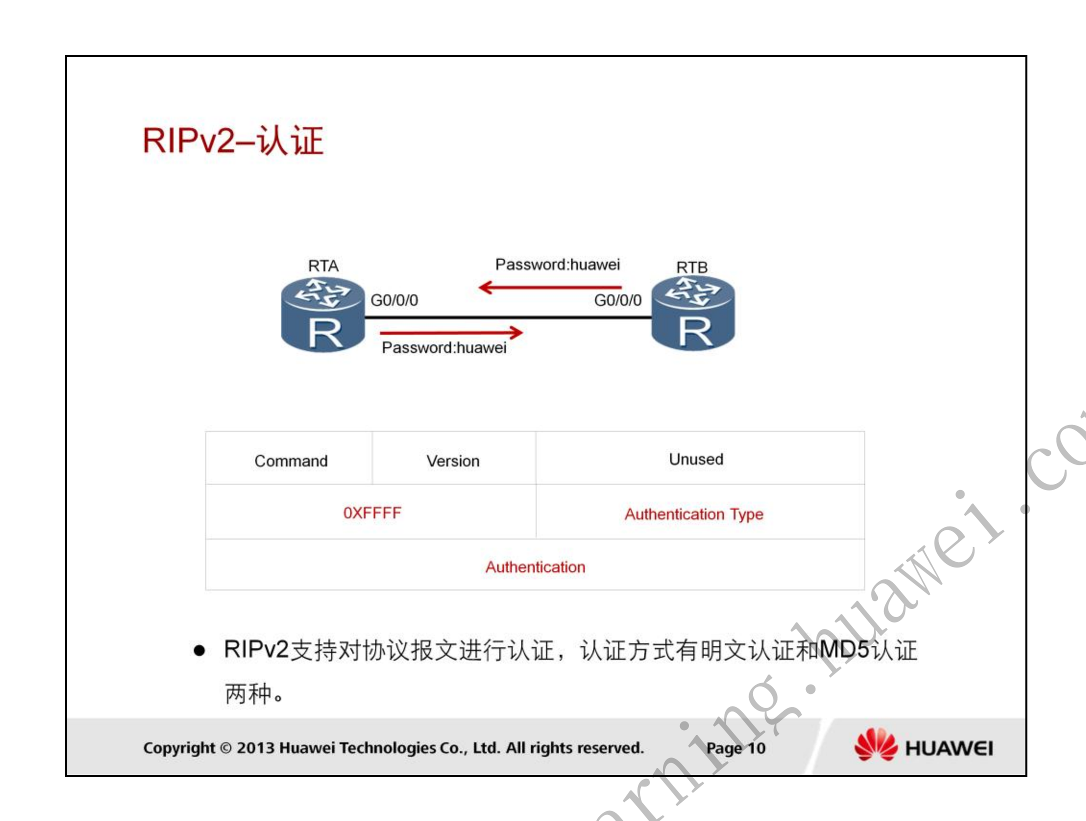
 Page 9 HUAWEI

RIPv2在RIPv1基础上进行了扩展，但RIPv2的报文格式仍然同RIPv1类似。

其中不同的字段如下所示：

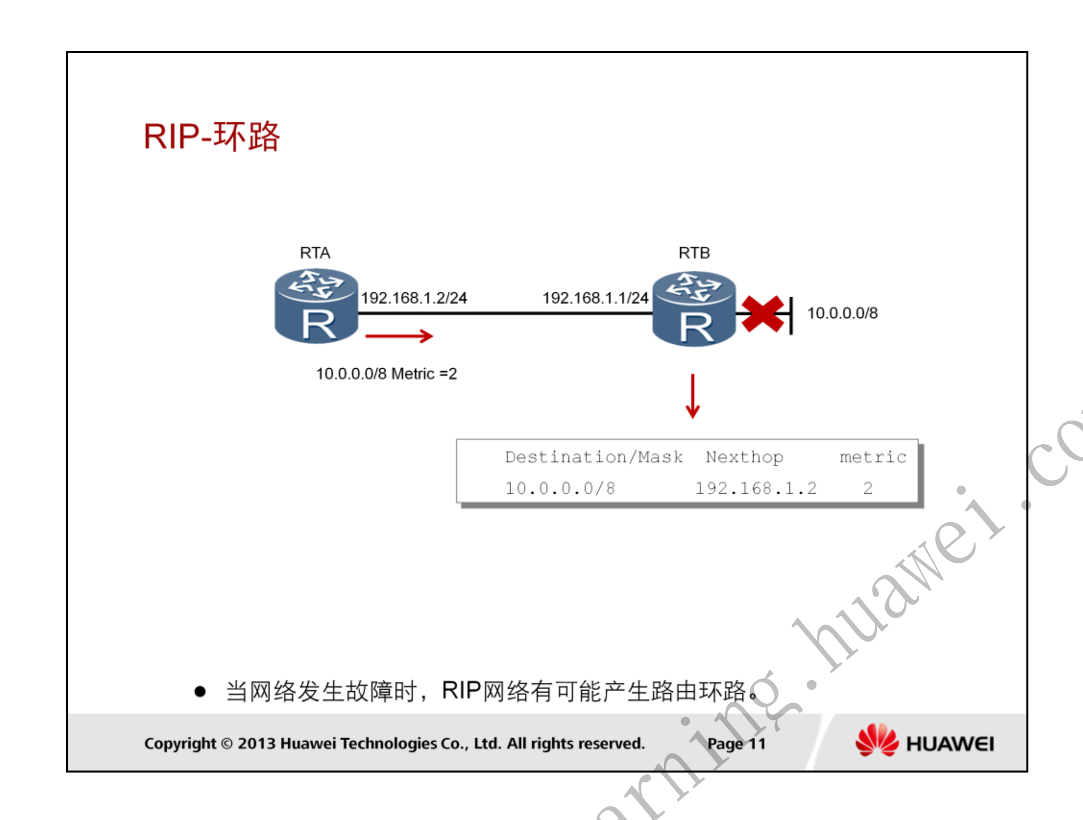
1. AFI:地址族标识除了表示支持的协议类型外，还可以用来描述认证信息。
2. Route tag：用于标记外部路由。
3. Subnet Mask：指定IP地址的子网掩码，定义IP地址的网络或子网部分。
4. Next Hop：指定通往目的地址的下一跳IP地址。

RFC1723对RIPv1和RIPv2的兼容性问题进行了分析和讨论，这里不再进行描述。



RIPv2的认证功能是一种过滤恶意路由信息的方法，该方法根据key值来检查从有效对端设备接收到的报文。这个key值是每个接口上都可以配置的一个明文密码串，相应的认证类型（Authentication Type）的值为2。

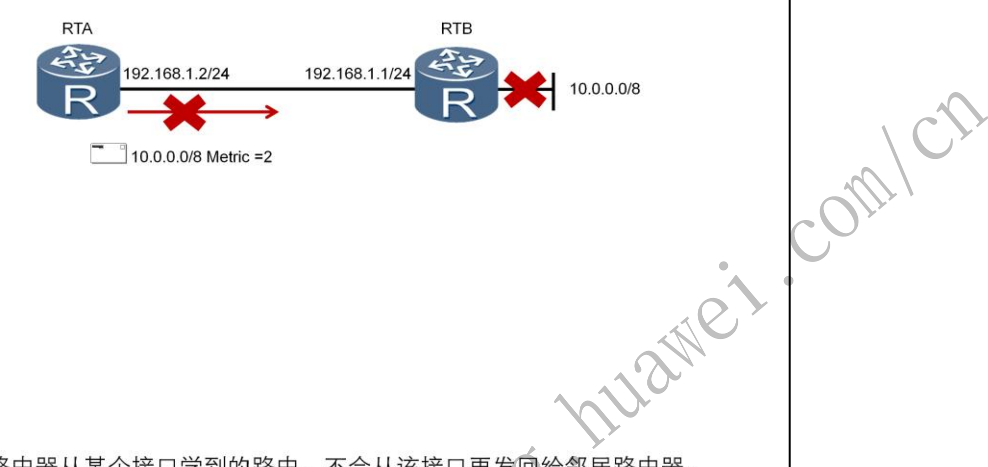
早期的RIPv2只支持简单明文认证，安全性低，因为明文认证密码串可以很轻易地截获。随着对RIP安全性的需求越来越高，RIPv2引入了加密认证功能，开始是通过支持MD5认证（RFC 2082）来实现，后来通过支持HMAC-SHA-1认证（RFC 2082）进一步增强了安全性。华为AR2200系列路由器能够支持以上提到的所有认证方式，本示例只介绍了简单明文认证原理。



本示例介绍了RIP网络上路由环路的形成。

如图所示，RIP网络正常运行时，RTA会通过RTB学习到10.0.0.0/8网络的路由，度量值为1。一旦路由器RTB的直连网络10.0.0.0/8产生故障，RTB会立即检测到该故障，并认为该路由不可达。此时，RTA还没有收到该路由不可达的信息，于是会继续向RTB发送度量值为2的通往10.0.0.0/8的路由信息。RTB会学习此路由信息，认为可以通过RTA到达10.0.0.0/8网络。此后，RTB发送的更新路由表，又会导致RTA路由表的更新，RTA会新增一条度量值为3的10.0.0.0/8网络路由表项，从而形成路由环路。这个过程会持续下去，直到度量值为16。

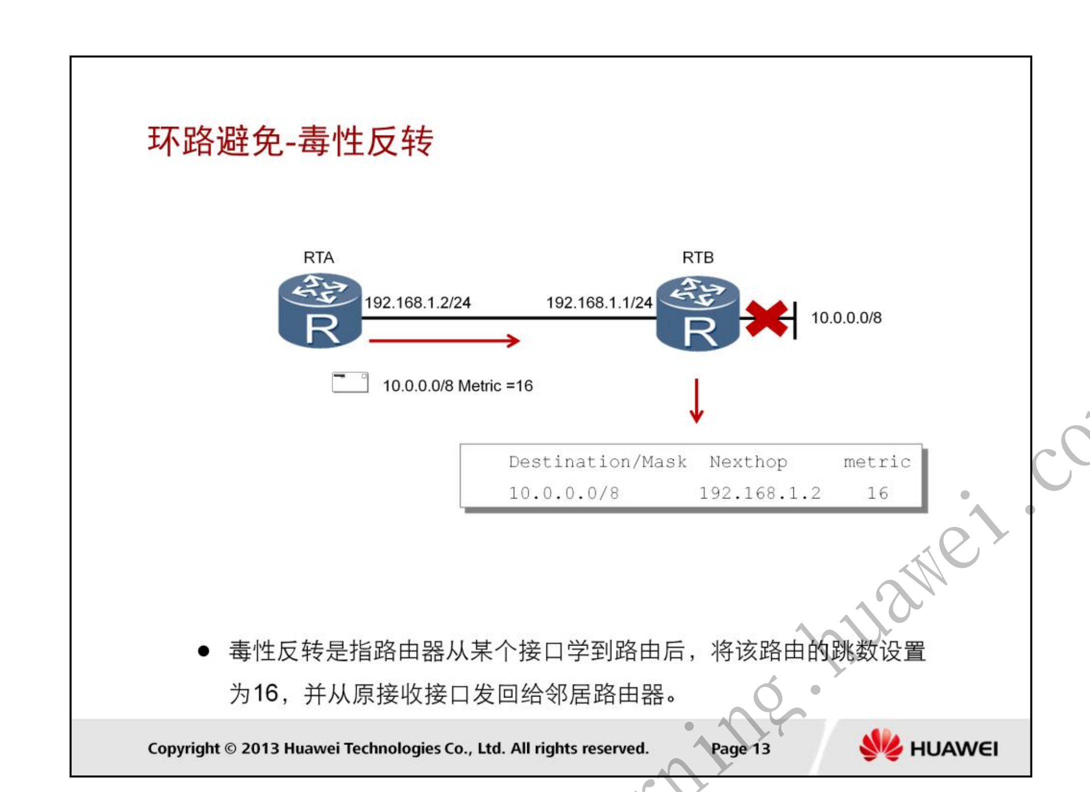
环路避免-水平分割



- 路由器从某个接口学到的路由，不会从该接口再发回给邻居路由器。

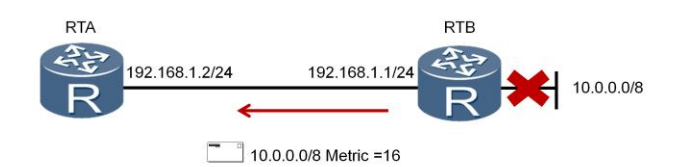
 Page 12 

RIP路由协议引入了很多机制来解决环路问题，除了之前介绍的最大跳数，还有水平分割机制。水平分割的原理是，路由器从某个接口学习到的路由，不会再从该接口发出去。也就是说，RTA从RTB学习到的10.0.0.0/8网络的路由不会再从RTA的接收接口重新通告给RTB，由此避免了路由环路的产生。



RIP的防环机制中还包括毒性反转，毒性反转机制的实现可以使错误路由立即超时。配置了毒性反转之后，RIP从某个接口学习到路由之后，发回给邻居路由器时会将该路由的跳数设置为16。利用这种方式，可以清除对方路由表中的无用路由。本示例中，RTB向RTA通告了度量值为1的10.0.0.0/8路由，RTA在通告给RTB时将该路由度量值设为16。如果10.0.0.0/8网络发生故障，RTB便不会认为可以通过RTA到达10.0.0.0/8网络，因此就可以避免路由环路的产生。

环路避免-触发更新



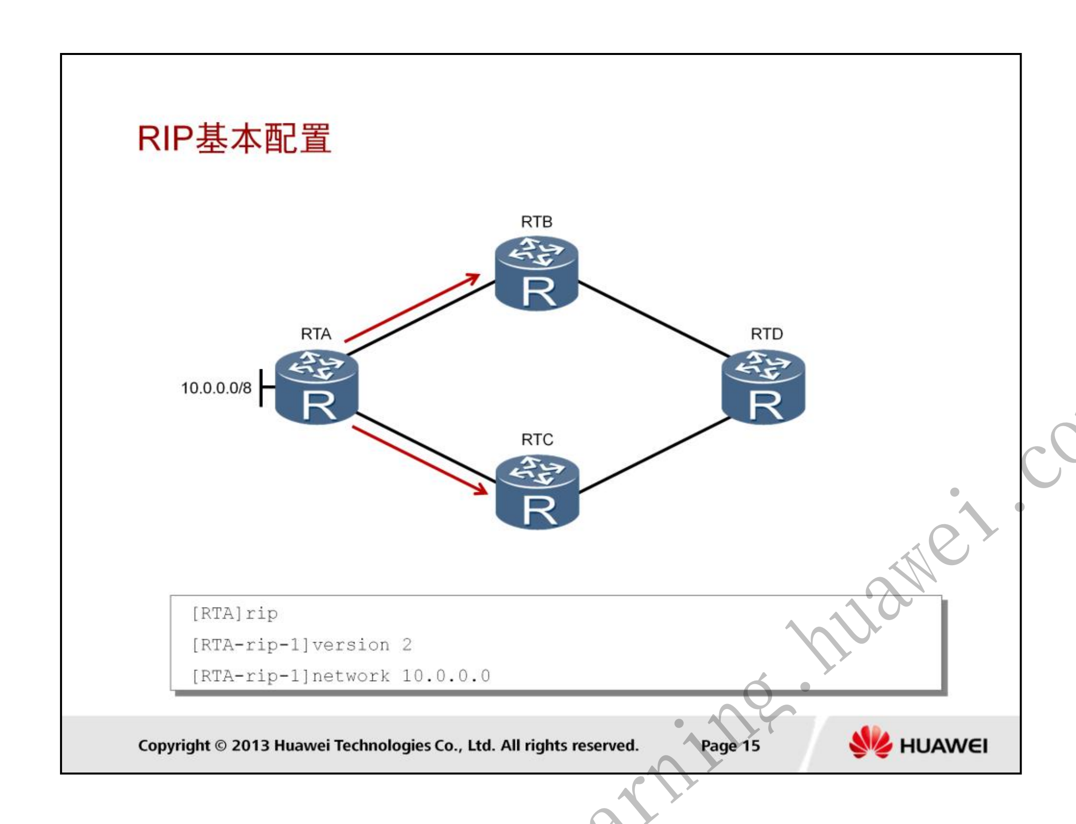


- 触发更新是指当路由信息发生变化时，立即向邻居设备发送触发更新报文。

 Page 14 HUAWEI

缺省情况下，一台RIP路由器每30秒会发送一次路由表更新给邻居路由器。

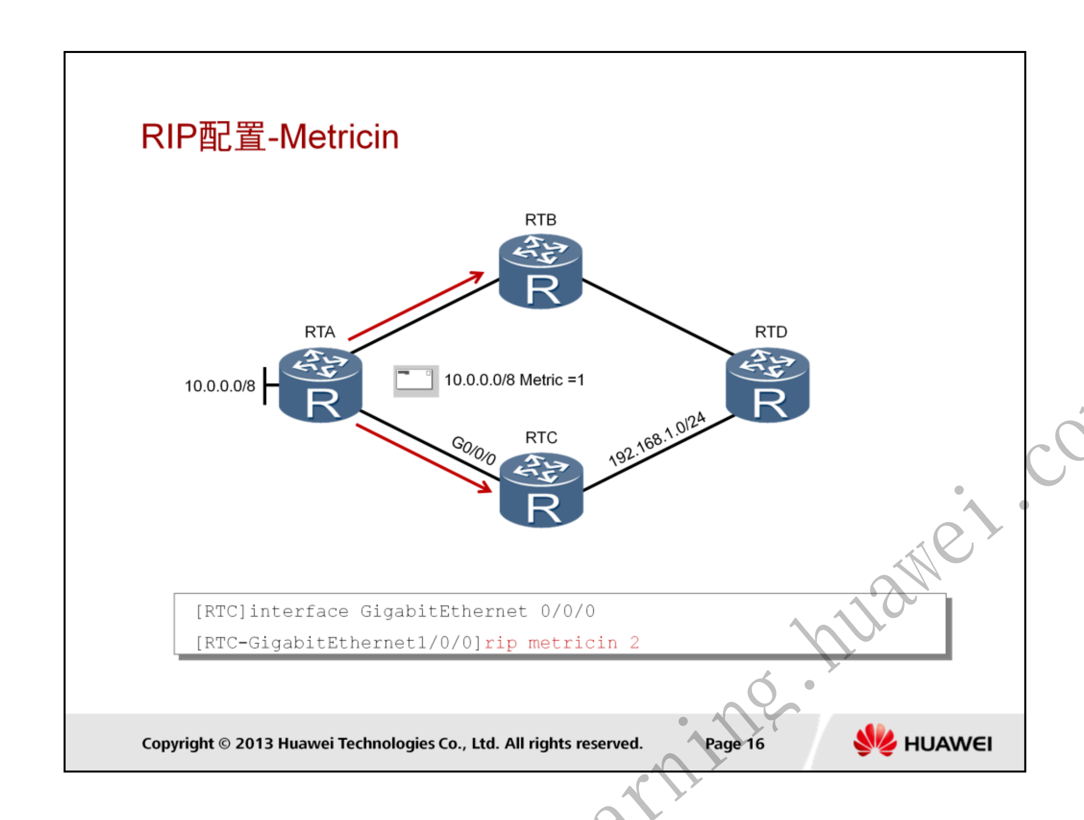
当本地路由信息发生变化时，触发更新功能允许路由器立即发送触发更新报文给邻居路由器，来通知路由信息更新，而不需要等待更新定时器超时，从而加速了网络收敛。



rip [*process-id*]命令用来使能RIP进程。该命令中，*process-id*指定了RIP进程ID。如果未指定*process-id*，命令将使用1作为缺省进程ID。

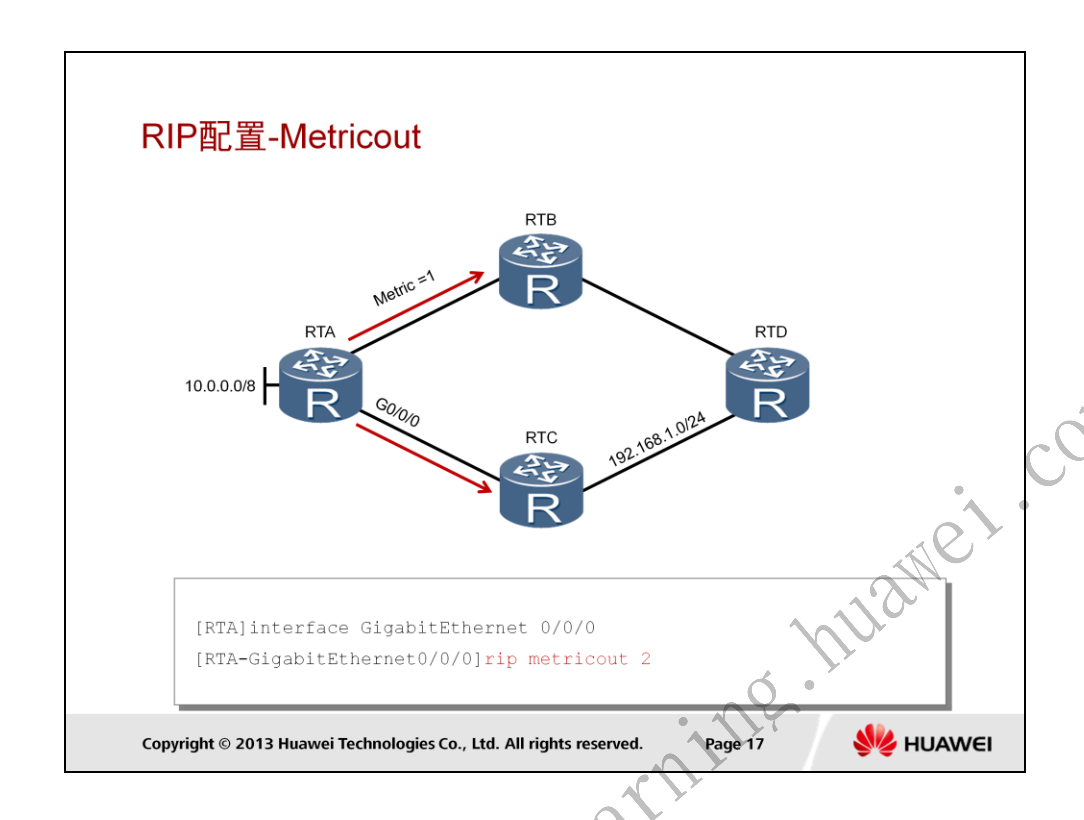
命令**version 2**可用于使能RIPv2以支持扩展能力，比如支持VLSM、认证等。

network *<network-address>*命令可用于在RIP中通告网络，*network-address*必须是一个自然网段的地址。只有处于此网络中的接口，才能进行RIP报文的接收和发送。



在RIP网络中，命令**rip metricin** <*metric value*>用于修改接口上应用的度量值（注意：该命令所指定的度量值会与当前路由的度量值相加）。当路由器的一个接口收到路由时，路由器会首先将接口的附加度量值增加到该路由上，然后将路由加入路由表中。

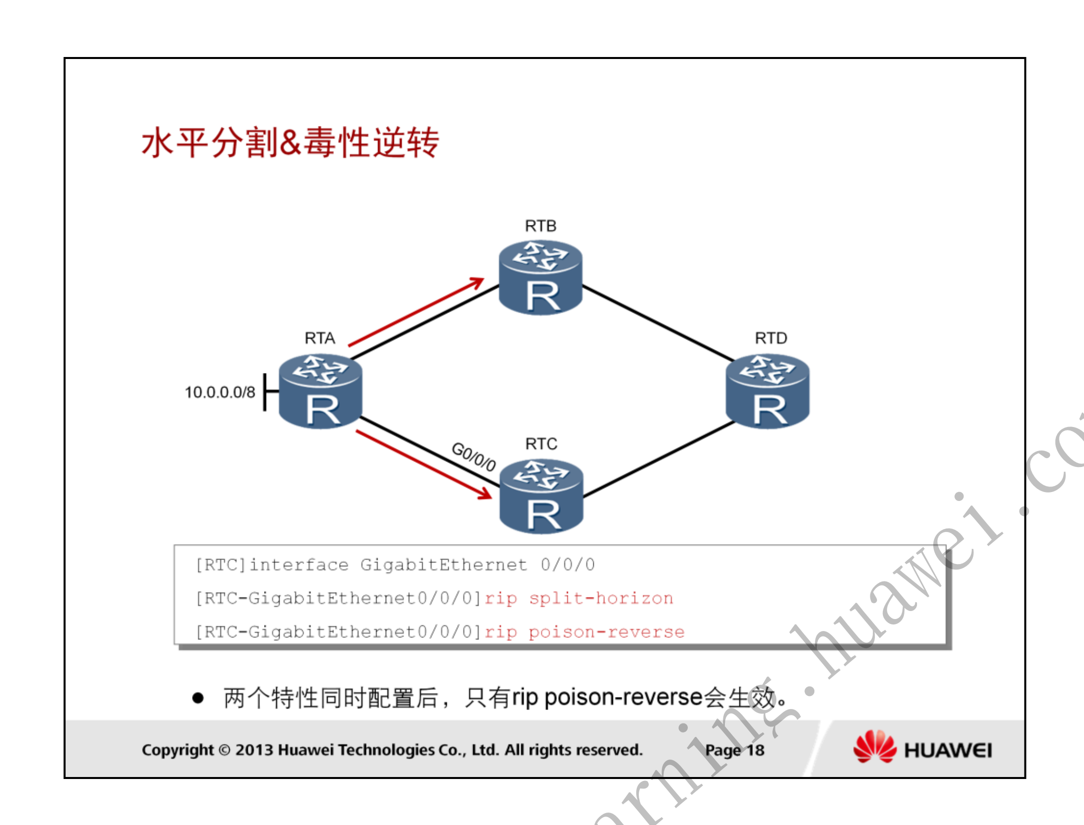
本示例中，RTA发送的10.0.0.0/8路由条目的度量值为1，由于在RTC的GigabitEthernet0/0/0接口上配置了rip metricin 2，所以当路由到达RTC的接口时，RTC会将该路由条目的度量值加2，最后该路由的度量值为3。



命令**rip metricout**用于路由器在通告RIP路由时修改路由的度量值。

一般情况下，在将路由表项转发到下一跳之前，RIP会将度量值加1。如果配置了**rip metricout**命令，则只应用命令中配置的度量值。即，当路由器发布一条路由时，此命令配置的度量值会在发布该路由之前附加在这条路由上，但本地路由表中的度量值不会发生改变。

在本示例中，缺省情况下，RTA发送的10.0.0.0/8路由条目的度量值为1。但是，由于在RTA的GigabitEthernet0/0/0接口上配置了**rip metricout 2**，所以RTA会将该路由条目的度量值设置为2，然后发送给RTC。



水平分割和毒性反转都是基于每个接口来配置的。缺省情况下，每个接口都启用了rip split-horizon命令（NBMA网络除外）以防止路由环路。华为ARG3系列路由器不支持同时配置水平分割和毒性反转，因此当一个接口上同时配置了水平分割和毒性反转时，只有毒性反转生效。

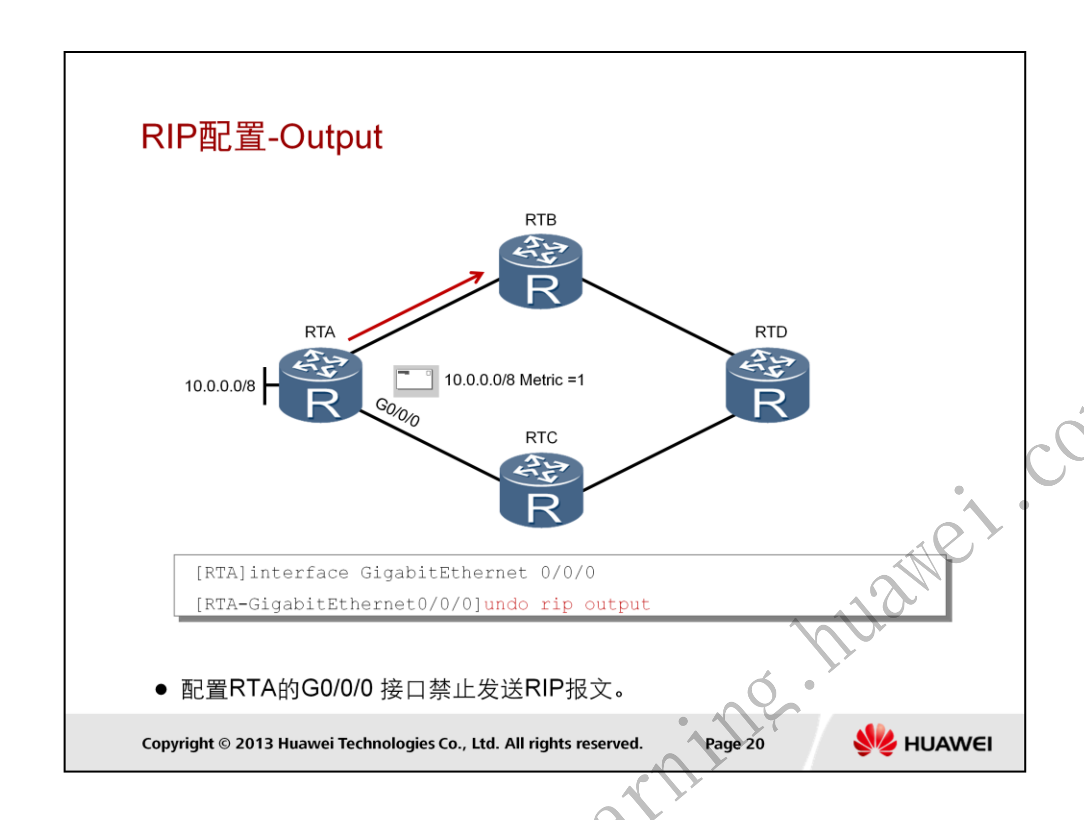
配置验证

```
[RTC] display rip 1 interface GigabitEthernet0/0/0 verbose
 GigabitEthernet0/0/0(192.168.1.2)
  State            : UP          MTU    : 500
  Metricin         : 2
  Metricout        : 1
  Input            : Enabled     Output : Enabled
  Protocol         : RIPv2 Multicast
  Send version     : RIPv2 Multicast Packets
  Receive version  : RIPv2 Multicast and Broadcast Packets
  Poison-reverse                : Enabled
  Split-Horizon                 : Enabled
  Authentication type           : None
  Replay Protection             : Disabled
```

- 虽然两者都显示“Enabled”，但只有“Poison-reverse”生效。

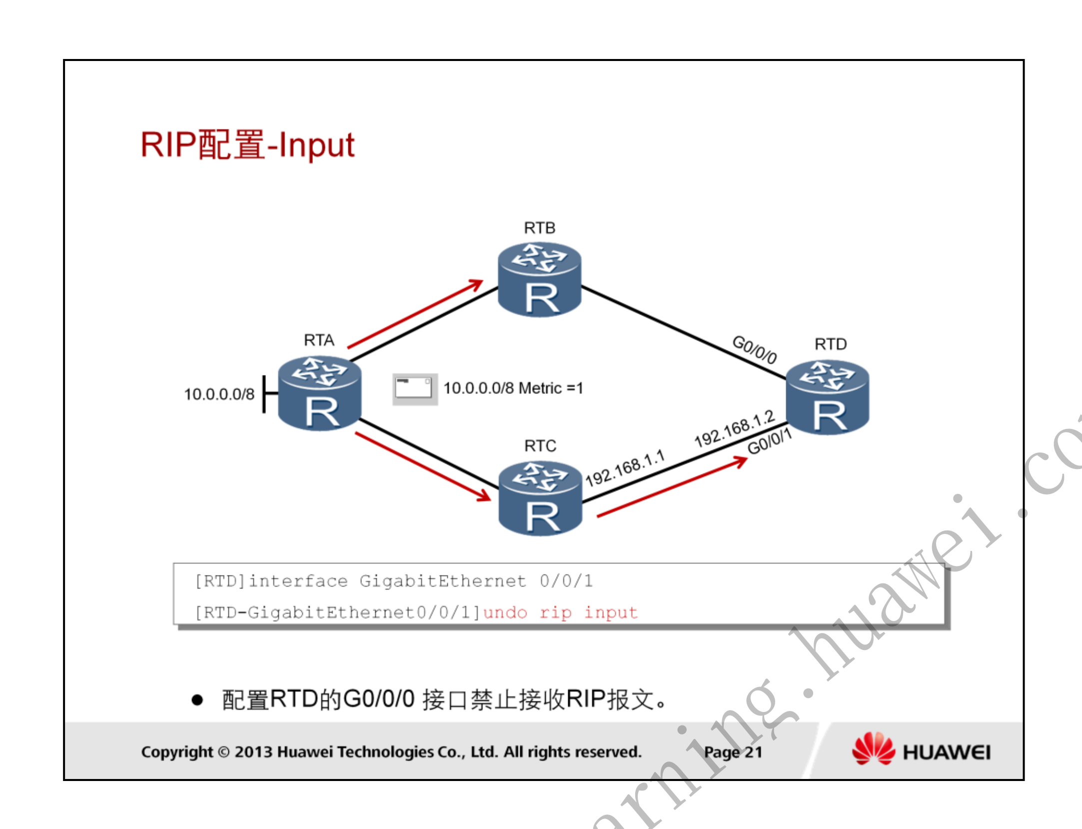


命令**display rip** *<process_id>* **interface** *<interface>* **verbose**用来确认路由器接口的RIP配置。命令回显中会显示相关RIP参数，包括RIP版本以及接口上是否应用了水平分割和毒性反转。此例中显示RTC的GigabitEthernet0/0/0接口配置了RIPv2，metricin为2，还启用了水平分割和毒性反转的功能。



命令**rip output**用于配置允许一个接口发送RIP更新消息。如果想要禁止指定接口发送RIP更新消息，可以在接口上运行命令**undo rip output**。缺省情况下，ARG3系列路由器允许接口发送RIP报文。

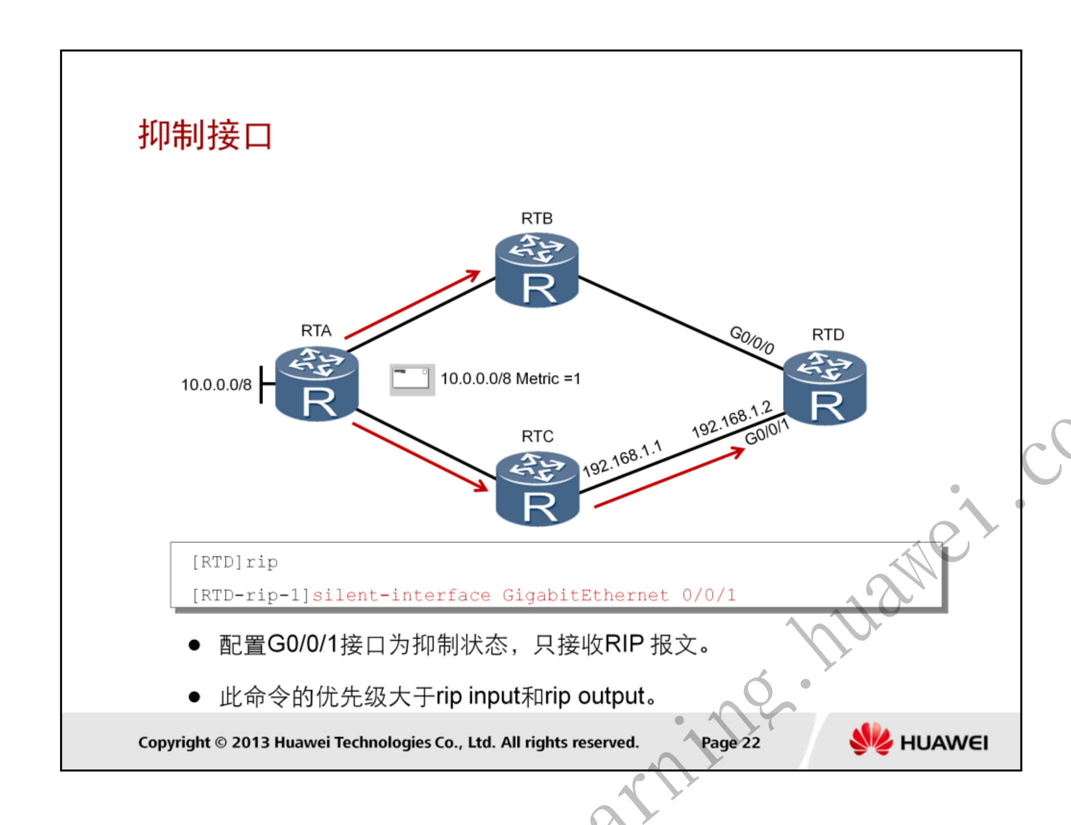
企业网络中，可以通过运行命令**undo rip output**来防止连接外网的接口发布内部路由。



rip input命令用来配置允许指定接口接收RIP报文。

undo rip input命令用来禁止指定接口接收RIP报文。运行命令**undo rip input**之后，该接口所收到的RIP报文会被立即丢弃。

缺省情况下，接口可以接收RIP报文。



silent-interface命令用来抑制接口，使其只接收RIP报文，更新自己的路由表，但不发送RIP报文。

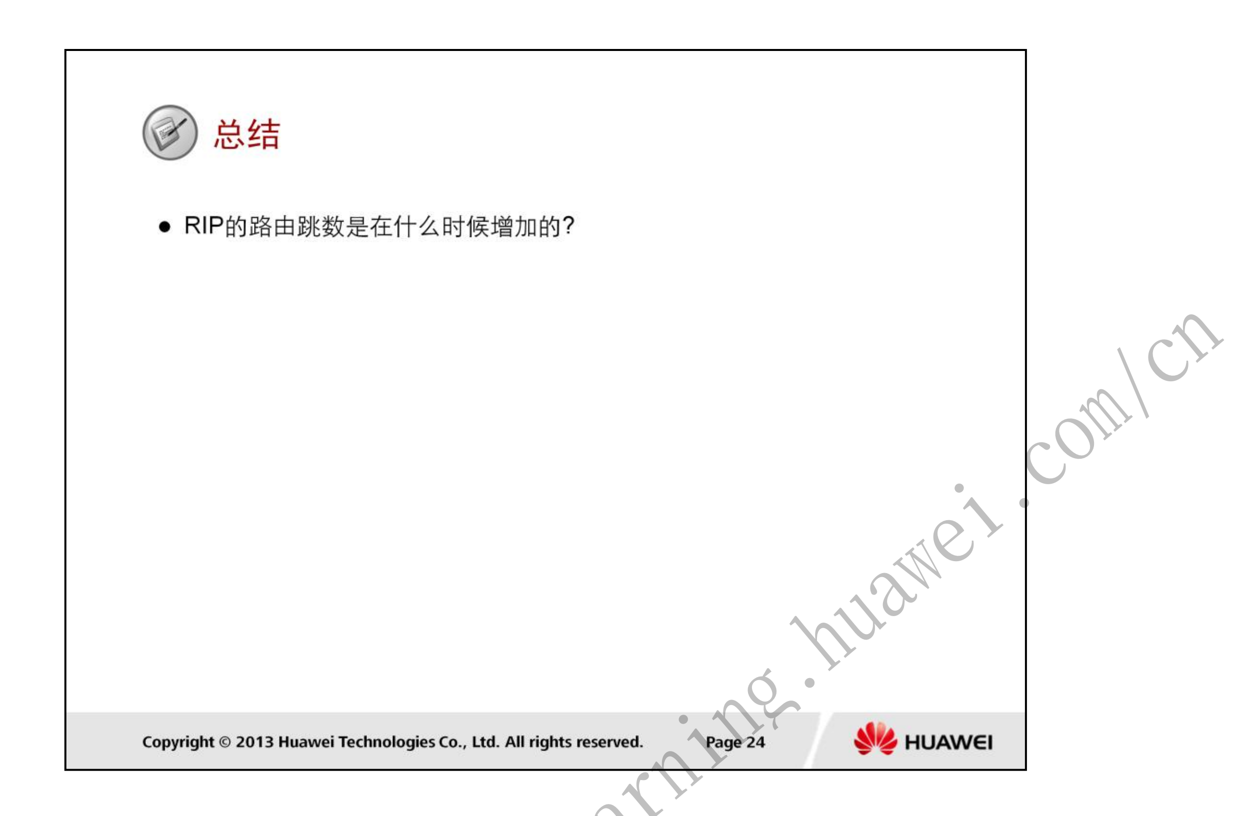
命令**silent-interface**比命令**rip input**和**rip output**的优先级更高。命令**silent-interface all**表示抑制所有接口，此命令优先级最高，在配置该命令之后，所有接口都被抑制。

命令**silent-interface**通常会配置在NBMA网络上。在NBMA网络上，一些路由器需要接收RIP更新消息但是不需要广播或者组播路由器自身的路由更新，而是通过命令**peer** *<ip address>*与对端路由器建立关系。

配置验证

```
[RTD] display rip
Public VPN-instance
    RIP process : 1
       RIP version    : 2
       Preference     : 100
       Checkzero      : Enabled
       Default-cost   : 0
       Summary        : Enabled
       Host-route     : Enabled
       Maximum number of balanced paths : 8
       Update time    : 30 sec              Age time : 180 sec
       Garbage-collect time : 120 sec
       Graceful restart  : Disabled
       BFD               : Disabled
       Silent-interfaces : GigabitEthernet0/0/1
```


命令**display rip**可以比较全面地显示路由器上的RIP信息，包括全局参数以及部分接口参数。例如，该命令可以显示哪些接口上执行了**silent-interface**命令。



1. RIP的路由跳数是在路由器发出路由通告之前增加的。


谢谢
www.huawei.com


链路状态路由协议-OSPF
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

开放式最短路径优先OSPF（Open Shortest Path First）协议是IETF定义的一种基于链路状态的内部网关路由协议。

RIP是一种基于距离矢量算法的路由协议，存在着收敛慢、易产生路由环路、可扩展性差等问题，目前已逐渐被OSPF取代。

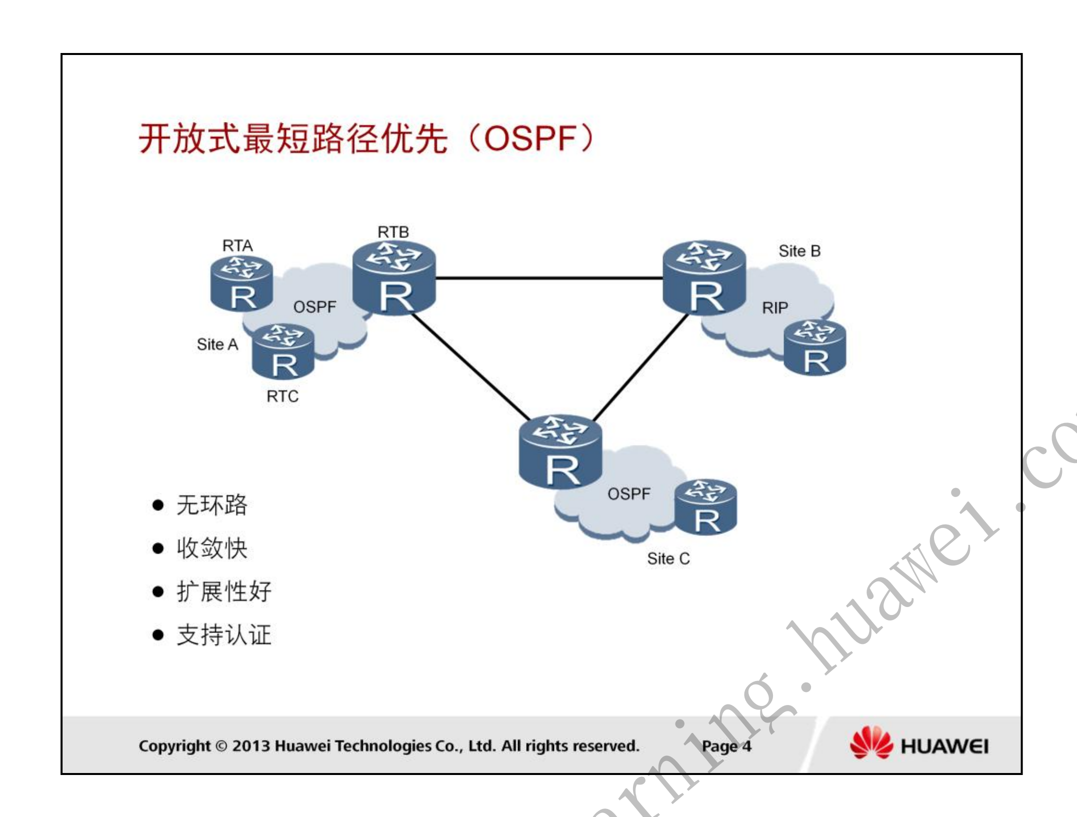




学习目标

学完本课程后，您应该能：

- 掌握OSPF的工作原理
- 掌握OSPF的基本配置

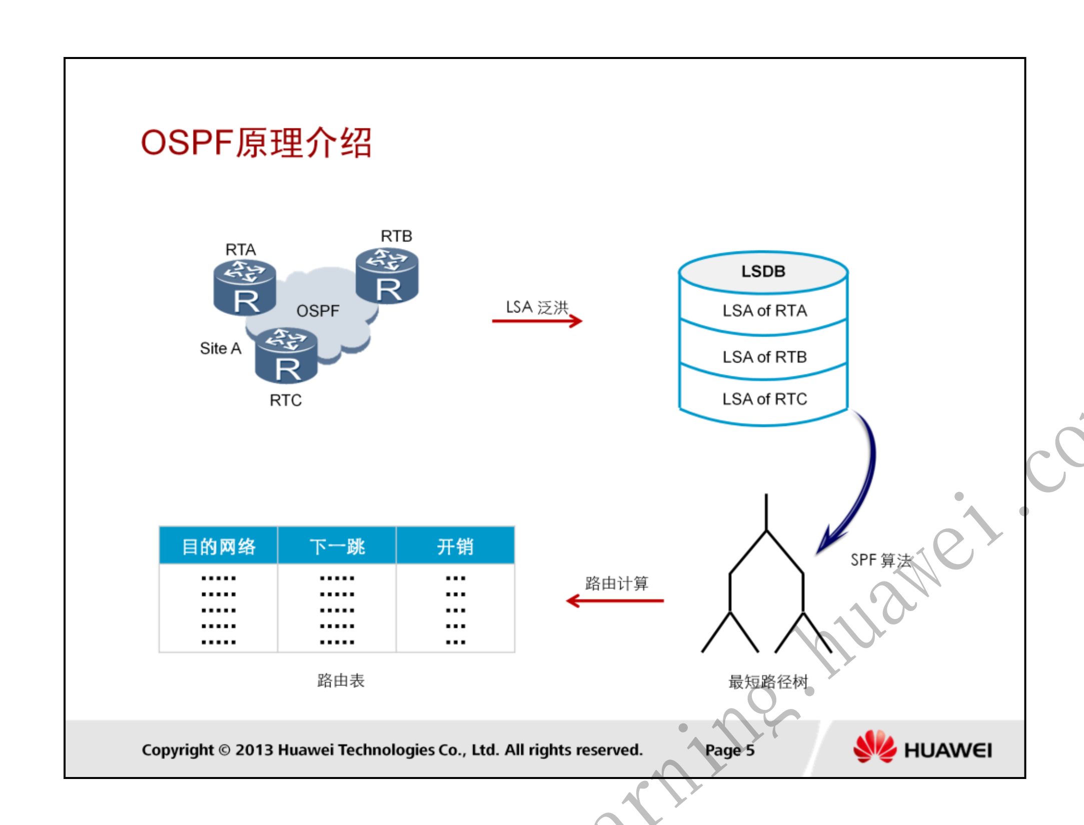


OSPF是一种基于链路状态的路由协议，它从设计上就保证了无路由环路。OSPF支持区域的划分，区域内部的路由器使用SPF最短路径算法保证了区域内部的无环路。OSPF还利用区域间的连接规则保证了区域之间无路由环路。

OSPF支持触发更新，能够快速检测并通告自治系统内的拓扑变化。

OSPF可以解决网络扩容带来的问题。当网络上路由器越来越多，路由信息流量急剧增长的时候，OSPF可以将每个自治系统划分为多个区域，并限制每个区域的范围。OSPF这种分区域的特点，使得OSPF特别适用于大中型网络。OSPF还可以同其他协议（比如多协议标记切换协议MPLS）同时运行来支持地理覆盖很广的网络。

OSPF可以提供认证功能。OSPF路由器之间的报文可以配置成必须经过认证才能进行交换。



OSPF要求每台运行OSPF的路由器都了解整个网络的链路状态信息，这样才能计算出到达目的地的最优路径。OSPF的收敛过程由链路状态公告LSA（Link State Advertisement）泛洪开始，LSA中包含了路由器已知的接口IP地址、掩码、开销和网络类型等信息。收到LSA的路由器都可以根据LSA提供的信息建立自己的链路状态数据库LSDB（Link State Database），并在LSDB的基础上使用SPF算法进行运算，建立起到达每个网络的最短路径树。最后，通过最短路径树得出到达目的网络的最优路由，并将其加入到IP路由表中。

OSPF报文

IP Header | OSPF Protocol Packet

- OSPF报文封装在IP报文中，协议号为89。
- OSPF报文类型有5种：
  - Hello 报文
  - DD（Database Description）报文
  - LSR（LSA Request）报文
  - LSU（LSA Update）报文
  - LSACK（Link State Acknowledgment）报文

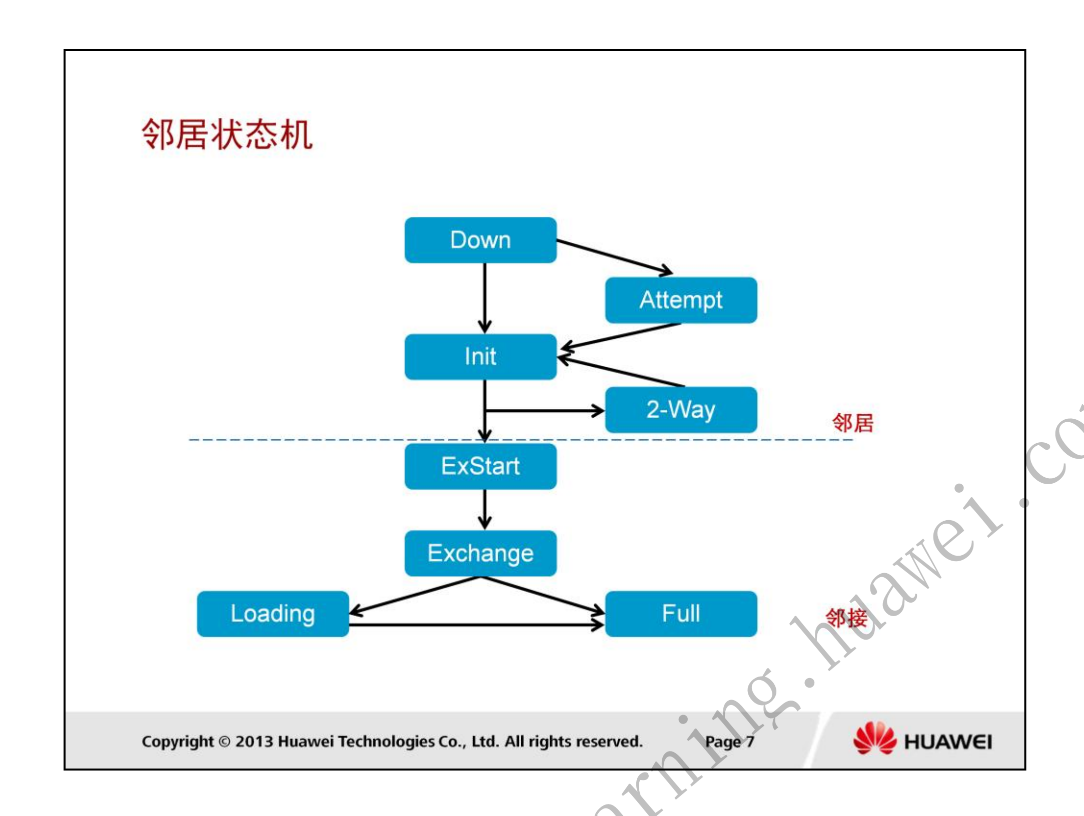
 Page 6



OSPF直接运行在IP协议之上，使用IP协议号89。

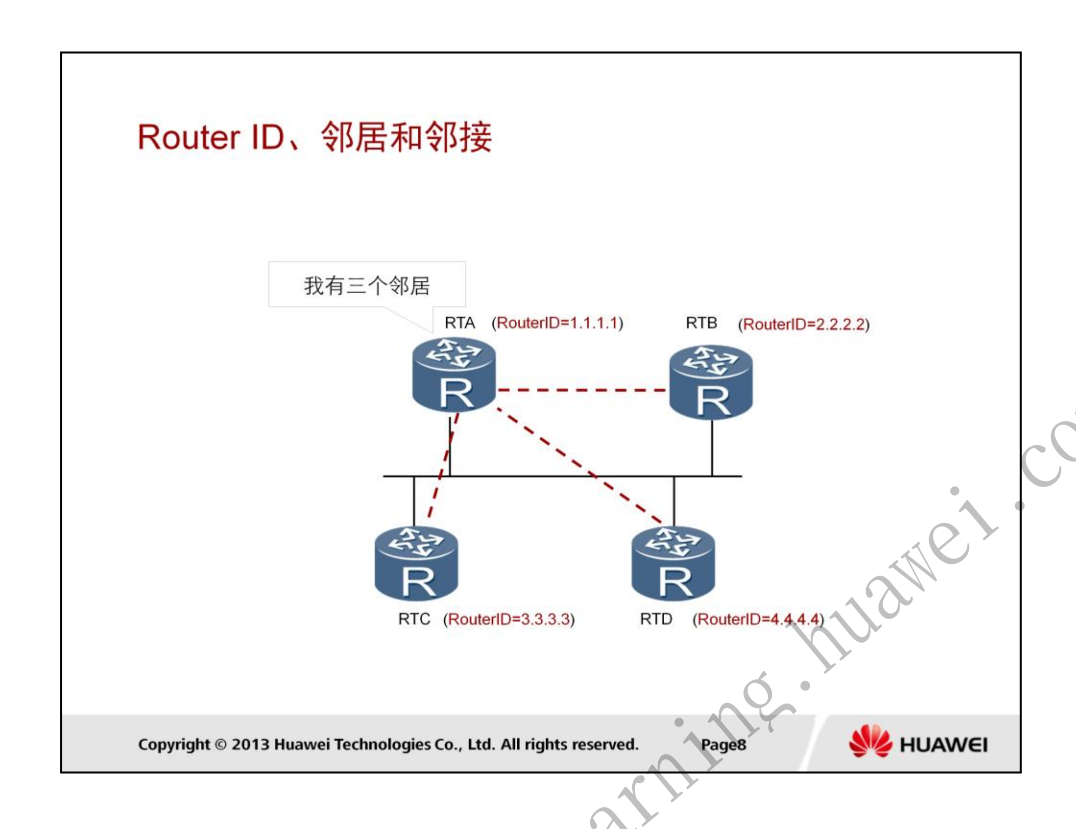
OSPF有五种报文类型，每种报文都使用相同的OSPF报文头。

1. Hello报文：最常用的一种报文，用于发现、维护邻居关系。并在广播和NBMA（None-Broadcast Multi-Access）类型的网络中选举指定路由器DR（Designated Router）和备份指定路由器BDR（Backup Designated Router）。
2. DD报文：两台路由器进行LSDB数据库同步时，用DD报文来描述自己的LSDB。DD报文的内容包括LSDB中每一条LSA的头部（LSA的头部可以唯一标识一条LSA）。LSA头部只占一条LSA的整个数据量的一小部分，所以，这样就可以减少路由器之间的协议报文流量。
3. LSR报文：两台路由器互相交换过DD报文之后，知道对端的路由器有哪些LSA是本地LSDB所缺少的，这时需要发送LSR报文向对方请求缺少的LSA，LSR只包含了所需要的LSA的摘要信息。
4. LSU报文：用来向对端路由器发送所需要的LSA。
5. LSACK报文：用来对接收到的LSU报文进行确认。



邻居和邻接关系建立的过程如下：

1. Down：这是邻居的初始状态，表示没有从邻居收到任何信息。
2. Attempt：此状态只在NBMA网络上存在，表示没有收到邻居的任何信息，但是已经周期性的向邻居发送报文，发送间隔为HelloInterval。如果RouterDeadInterval间隔内未收到邻居的Hello报文，则转为Down状态。
3. Init：在此状态下，路由器已经从邻居收到了Hello报文，但是自己不在所收到的Hello报文的邻居列表中，尚未与邻居建立双向通信关系。
4. 2-Way：在此状态下，双向通信已经建立，但是没有与邻居建立邻接关系。这是建立邻接关系以前的最高级状态。
5. ExStart：这是形成邻接关系的第一个步骤，邻居状态变成此状态以后，路由器开始向邻居发送DD报文。主从关系是在此状态下形成的，初始DD序列号也是在此状态下决定的。在此状态下发送的DD报文不包含链路状态描述。
6. Exchange：此状态下路由器相互发送包含链路状态信息摘要的DD报文，描述本地LSDB的内容。
7. Loading：相互发送LSR报文请求LSA，发送LSU报文通告LSA。
8. Full：路由器的LSDB已经同步。



Router ID是一个32位的值，它唯一标识了一个自治系统内的路由器，可以为每台运行OSPF的路由器上可以手动配置一个Router ID，或者指定一个IP地址作为Router ID。如果设备存在多个逻辑接口地址，则路由器使用逻辑接口中最大的IP地址作为Router ID；如果没有配置逻辑接口，则路由器使用物理接口的最大IP地址作为Router ID。在为一台运行OSPF的路由器配置新的Router ID后，可以在路由器上通过重置OSPF进程来更新Router ID。通常建议手动配置Router ID，以防止Router ID因为接口地址的变化而改变。

运行OSPF的路由器之间需要交换链路状态信息和路由信息，在交换这些信息之前路由器之间首先需要建立邻接关系。

邻居（Neighbor）：

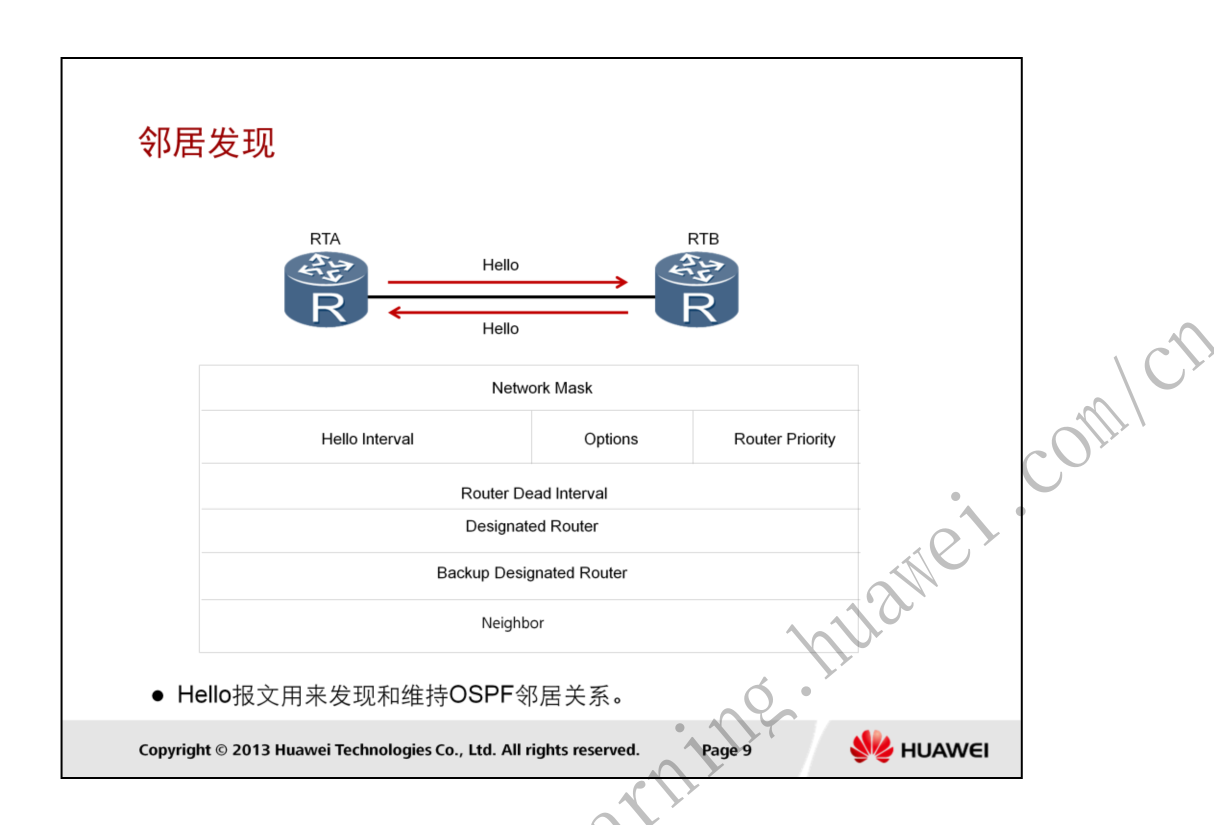
OSPF路由器启动后，便会通过OSPF接口向外发送Hello报文用于发现邻居。收到Hello报文的OSPF路由器会检查报文中所定义的一些参数，如果双方的参数一致，就会彼此形成邻居关系。

邻接（Adjacency）：

形成邻居关系的双方不一定都能形成邻接关系，这要根据网络类型而定。只有当双方成功交换DD报文，并能交换LSA之后，才形成真正意义上的邻接关系。

路由器在发送LSA之前必须先发现邻居并建立邻居关系。

本例中，RTA通过以太网连接了三个路由器，所以RTA有三个邻居，但不能说RTA有三邻接关系。



OSPF的邻居发现过程是基于Hello报文来实现的，Hello报文中的重要字段解释如下：

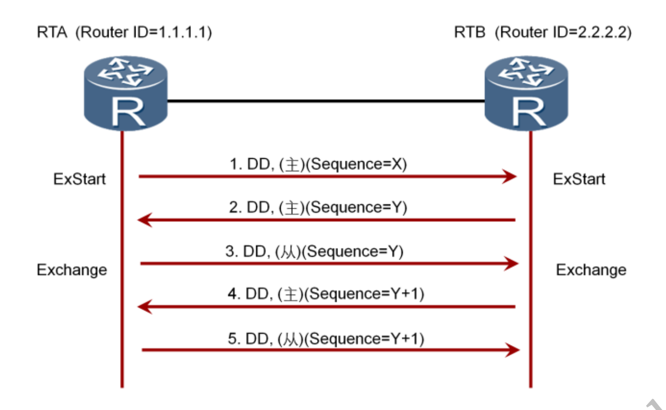
1. Network Mask：发送Hello报文的接口的网络掩码。
2. HelloInterval：发送Hello报文的时间间隔，单位为秒。
3. Options：标识发送此报文的OSPF路由器所支持的可选功能。具体的可选功能已超出这里的讨论范围。
4. Router Priority：发送Hello报文的接口的Router Priority，用于选举DR和BDR。
5. RouterDeadInterval：失效时间。如果在此时间内未收到邻居发来的Hello报文，则认为邻居失效；单位为秒，通常为四倍HelloInterval。
6. Designated Router：发送Hello报文的路由器所选举出的DR的IP地址。如果设置为0.0.0.0，表示未选举DR路由器。
7. Backup Designated Router：发送Hello报文的路由器所选举出的BDR的IP地址。如果设置为0.0.0.0，表示未选举BDR。
8. Neighbor：邻居的Router ID列表，表示本路由器已经从这些邻居收到了合法的Hello报文。

如果路由器发现所接收的合法Hello报文的邻居列表中有自己的Router ID，则认为已经和邻居建立了双向连接，表示邻居关系已经建立。

验证一个接收到的Hello报文是否合法包括：

1. 如果接收端口的网络类型是广播型，点到多点或者NBMA，所接收的Hello报文中Network Mask字段必须和接收端口的网络掩码一致，如果接收端口的网络类型为点到点类型或者是虚连接，则不检查Network Mask字段；
2. 所接收的Hello报文中Hello Interval字段必须和接收端口的配置一致；
3. 所接收的Hello报文中Router Dead Interval字段必须和接收端口的配置一致；
4. 所接收的Hello报文中Options字段中的E-bit（表示是否接收外部路由信息）必须和相关区域的配置一致。

数据库同步



- 路由器使用DD报文来进行主从路由器的选举和数据库摘要信息的交互。
- DD报文包含LSA的头部信息，用来描述LSDB的摘要信息。

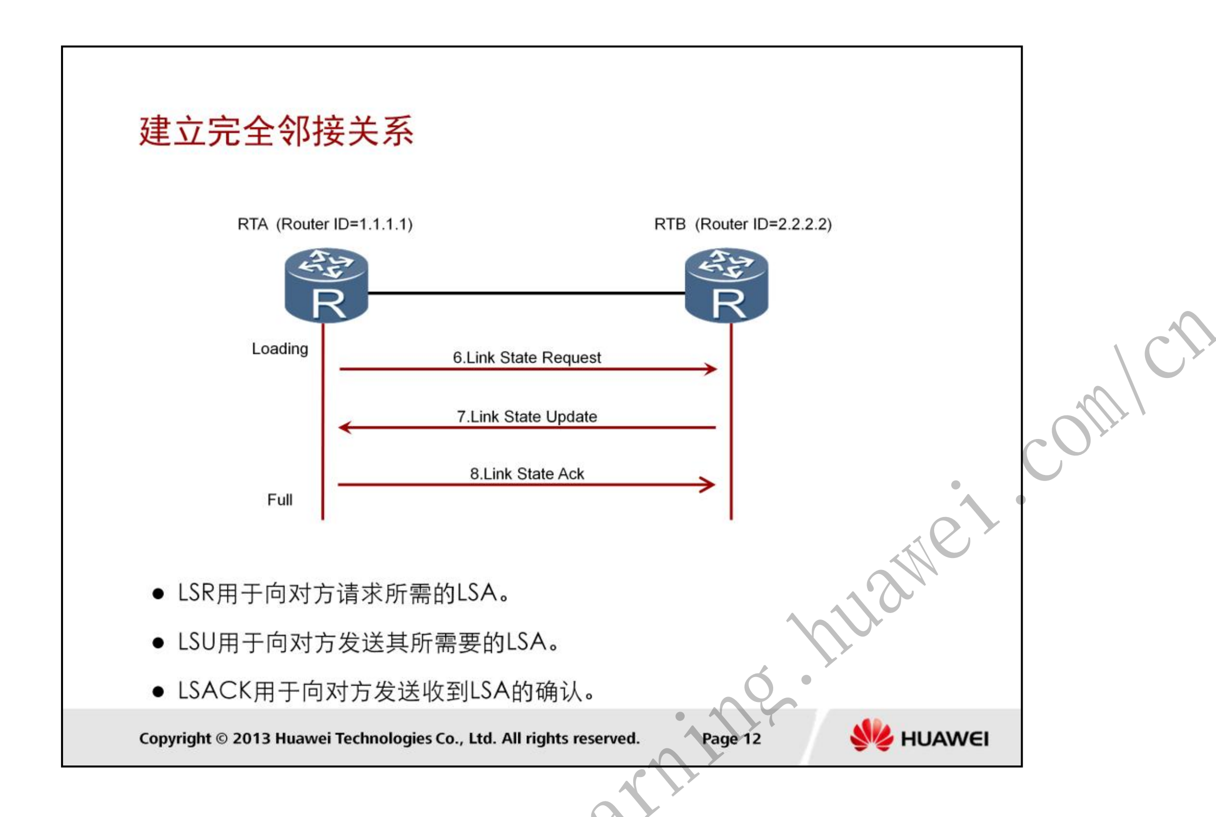
 Page 11



如图所示，路由器在建立完成邻居关系之后，便开始进行数据库同步，具体过程如下：

1. 邻居状态变为ExStart以后，RTA向RTB发送第一个DD报文，在这个报文中，DD序列号被设置为X（假设），RTA宣告自己为主路由器。
2. RTB也向RTA发送第一个DD报文，在这个报文中，DD序列号被设置为Y（假设）。RTB也宣告自己为主路由器。由于RTB的Router ID比RTA的大，所以RTB应当为真正的主路由器。
3. RTA发送一个新的DD报文，在这个新的报文中包含LSDB的摘要信息，序列号设置为RTB在步骤2里使用的序列号，因此RTB将邻居状态改变为Exchange。
4. 邻居状态变为Exchange以后，RTB发送一个新的DD报文，该报文中包含LSDB的描述信息，DD序列号设为Y+1（上次使用的序列号加1）。
5. 即使RTA不需要新的DD报文描述自己的LSDB，但是作为从路由器，RTA需要对主路由器RTB发送的每一个DD报文进行确认。所以，RTA向RTB发送一个内容为空的DD报文，序列号为Y+1。

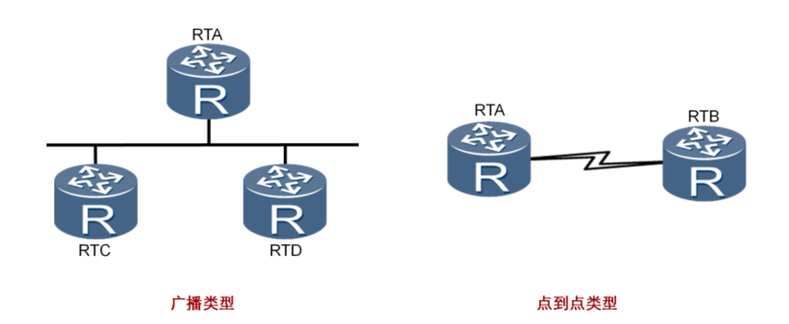
发送完最后一个DD报文之后，RTA将邻居状态改变为Loading；RTB收到最后一个DD报文之后，改变状态为Full（假设RTB的LSDB是最新最全的，不需要向RTA请求更新）。



6. 邻居状态变为Loading之后，RTA开始向RTB发送LSR报文，请求那些在Exchange状态下通过DD报文发现的，而且在本地LSDB中没有的链路状态信息。
7. RTB收到LSR报文之后，向RTA发送LSU报文，在LSU报文中，包含了那些被请求的链路状态的详细信息。RTA收到LSU报文之后，将邻居状态从Loading改变成Full。
8. RTA向RTB发送LSACK报文，用于对已接收LSA的确认。

此时，RTA和RTB之间的邻居状态变成Full，表示达到完全邻接状态。

OSPF支持的网络类型



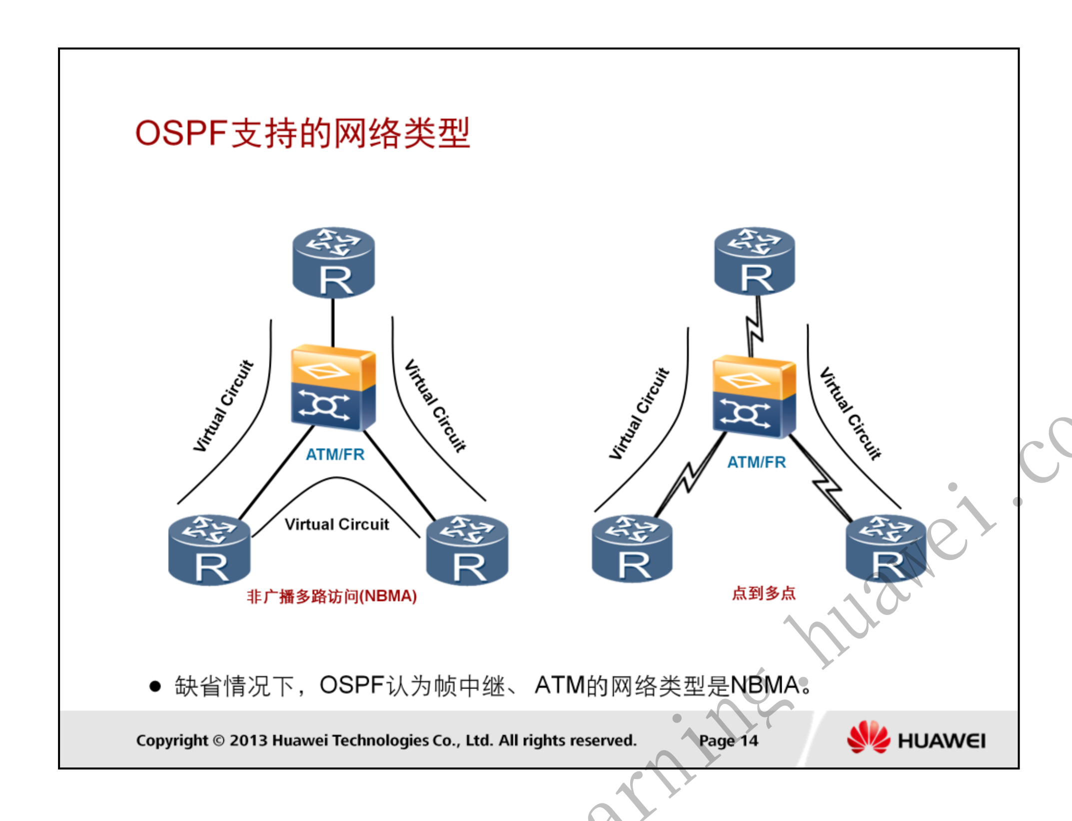
- 缺省情况下，OSPF认为以太网的网络类型是广播类型，PPP、HDLC的网络类型是点到点类型。

 Page 13 

OSPF定义了四种网络类型，分别是点到点网络，广播型网络，NBMA网络和点到多点网络。

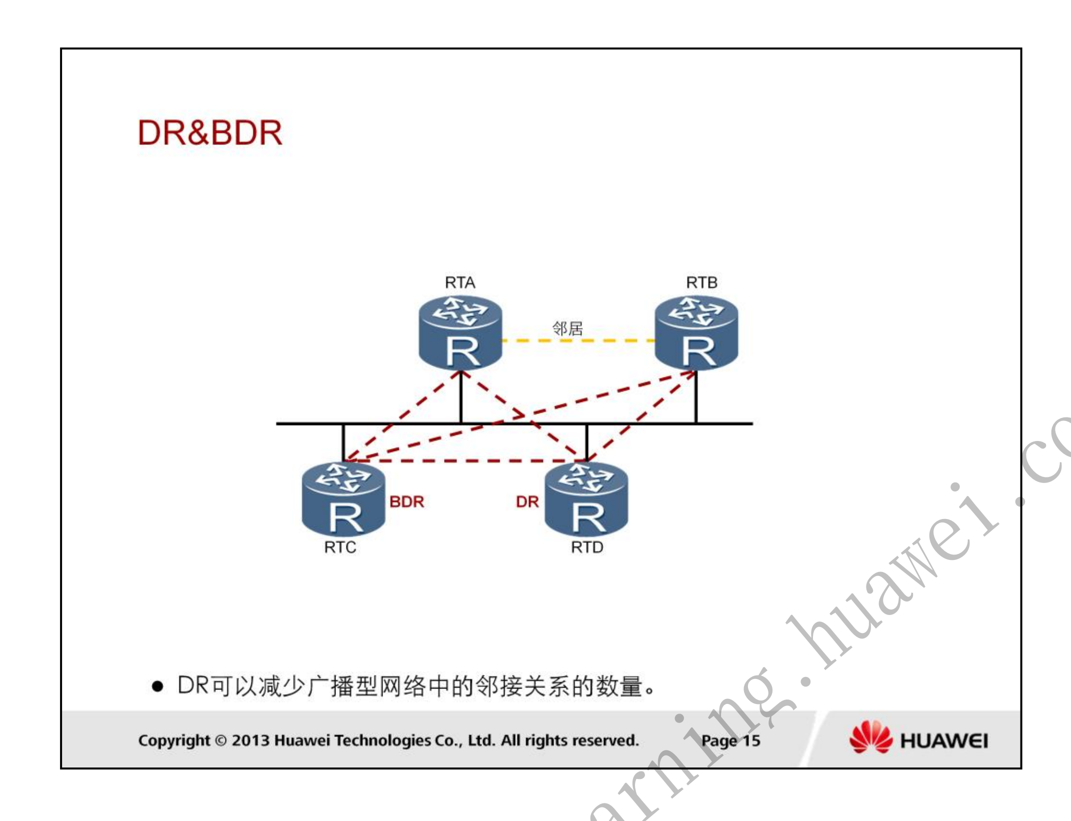
点到点网络是指只把两台路由器直接相连的网络。一个运行PPP的64K串行线路就是一个点到点网络的例子。

广播型网络是指支持两台以上路由器，并且具有广播能力的网络。一个含有三台路由器的以太网就是一个广播型网络的例子。



OSPF可以在不支持广播的多路访问网络上运行，此类网络包括在hub-spoke拓扑上运行的帧中继（FR）和异步传输模式（ATM）网络，这些网络的通信依赖于虚电路。OSPF定义了两种支持多路访问的网络类型：非广播多路访问网络（NBMA）和点到多点网络（Point To Multi-Points）。

1. NBMA:在NBMA网络上，OSPF模拟在广播型网络上的操作，但是每个路由器的邻居需要手动配置。NBMA方式要求网络中的路由器组成全连接。
2. P2MP:将整个网络看成是一组点到点网络。对于不能组成全连接的网络应当使用点到多点方式，例如只使用PVC的不完全连接的帧中继网络。

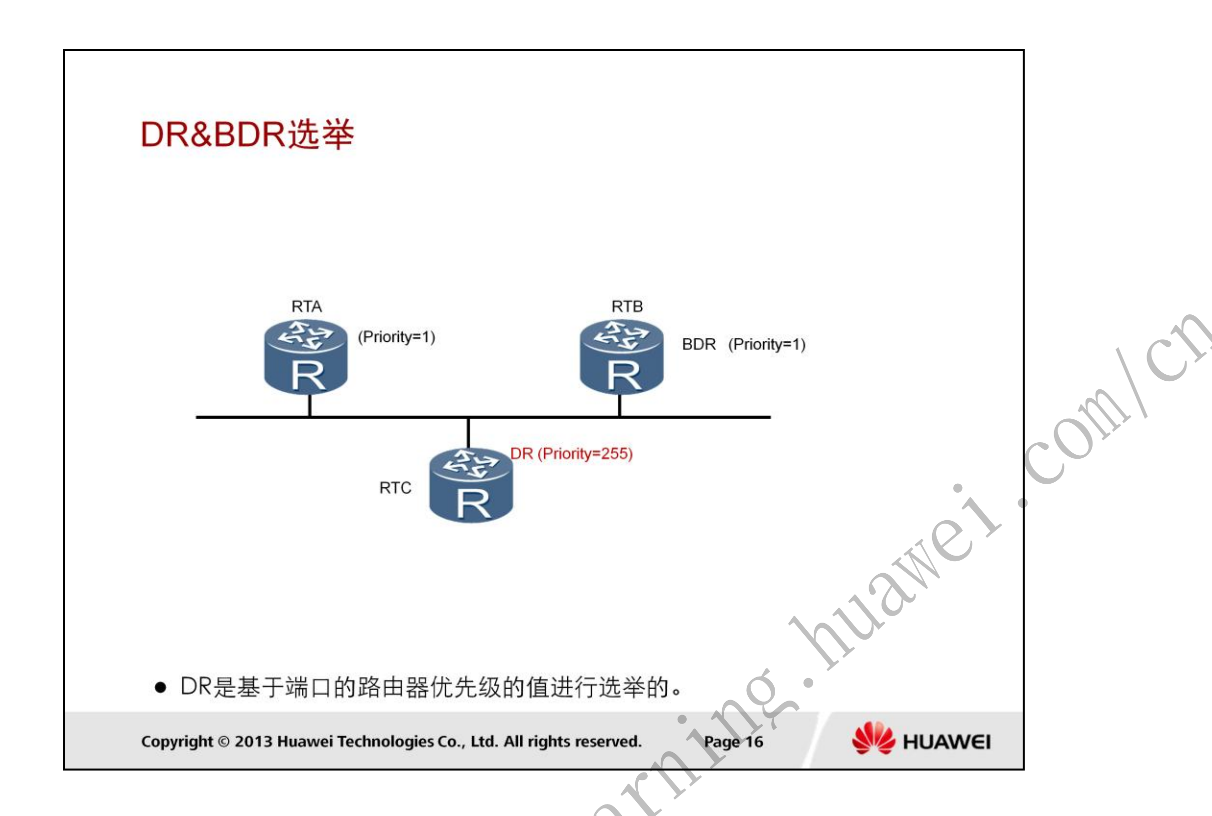


每一个含有至少两个路由器的广播型网络和NBMA网络都有一个DR和BDR。

DR和BDR可以减少邻接关系的数量，从而减少链路状态信息以及路由信息的交换次数，这样可以节省带宽，降低对路由器处理能力的压力。一个既不是DR也不是BDR的路由器只与DR和BDR形成邻接关系并交换链路状态信息以及路由信息，这样就大大减少了大型广播型网络和NBMA网络中的邻接关系数量。在没有DR的广播网络上，邻接关系的数量可以根据公式$n(n-1)/2$计算出，n代表参与OSPF的路由器接口的数量。在本例中，所有路由器之间有6个邻接关系。当指定了DR后，所有的路由器都与DR建立起邻接关系，DR成为该广播网络上的中心点。

BDR在DR发生故障时接管业务，一个广播网络上所有路由器都必须同BDR建立邻接关系。本例中使用DR和BDR将邻接关系从6减少到了5，RTA和RTB都只需要同DR和BDR建立邻接关系，RTA和RTB之间建立的是邻居关系。

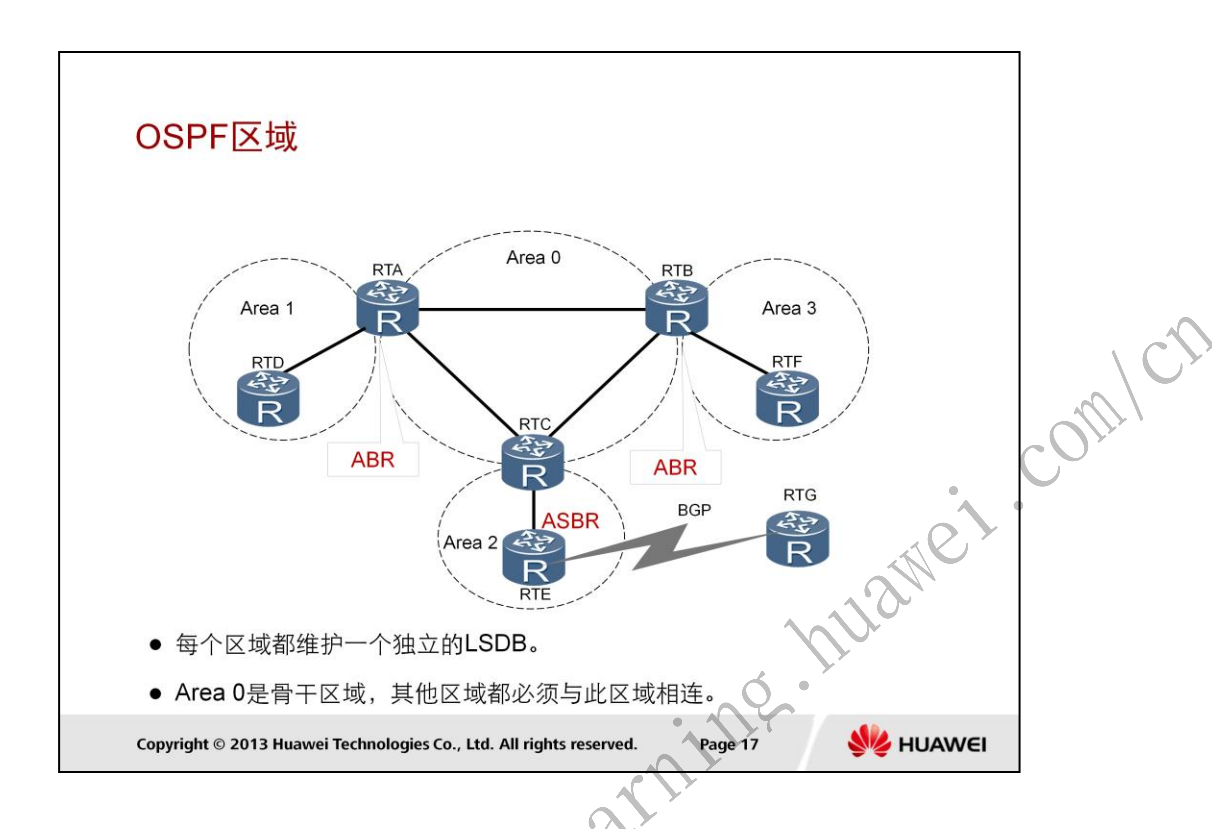
此例中，邻接关系数量的减少效果并不明显。但是，当网络上部署了大量路由器时，比如100台，那么情况就大不一样了。



在邻居发现完成之后，路由器会根据网段类型进行DR选举。在广播和NBMA网络上，路由器会根据参与选举的每个接口的优先级进行DR选举。优先级取值范围为0-255，值越高越优先。缺省情况下，接口优先级为1。如果一个接口优先级为0，那么该接口将不会参与DR或者BDR的选举。如果优先级相同时，则比较Router ID，值越大越优先被选举为DR。

为了给DR做备份，每个广播和NBMA网络上还要选举一个BDR。BDR也会与网络上所有的路由器建立邻接关系。

为了维护网络上邻接关系的稳定性，如果网络中已经存在DR和BDR，则新添加进该网络的路由器不会成为DR和BDR，不管该路由器的Router Priority是否最大。如果当前DR发生故障，则当前BDR自动成为新的DR，网络中重新选举BDR；如果当前BDR发生故障，则DR不变，重新选举BDR。这种选举机制的目的是为了保持邻接关系的稳定，使拓扑结构的改变对邻接关系的影响尽量小。



OSPF支持将一组网段组合在一起，这样的一个组合称为一个区域。

划分OSPF区域可以缩小路由器的LSDB规模，减少网络流量。

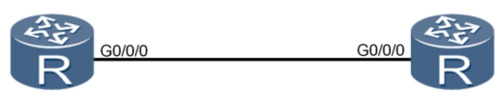
区域内的详细拓扑信息不向其他区域发送，区域间传递的是抽象的路由信息，而不是详细的描述拓扑结构的链路状态信息。每个区域都有自己的LSDB，不同区域的LSDB是不同的。路由器会为每一个自己所连接到的区域维护一个单独的LSDB。由于详细链路状态信息不会被发布到区域以外，因此LSDB的规模大大缩小了。

Area 0为骨干区域，为了避免区域间路由环路，非骨干区域之间不允许直接相互发布路由信息。因此，每个区域都必须连接到骨干区域。

运行在区域之间的路由器叫做区域边界路由器ABR（Area Boundary Router），它包含所有相连区域的LSDB。自治系统边界路由器ASBR（Autonomous System Boundary Router）是指和其他AS中的路由器交换路由信息的路由器，这种路由器会向整个AS通告AS外部路由信息。

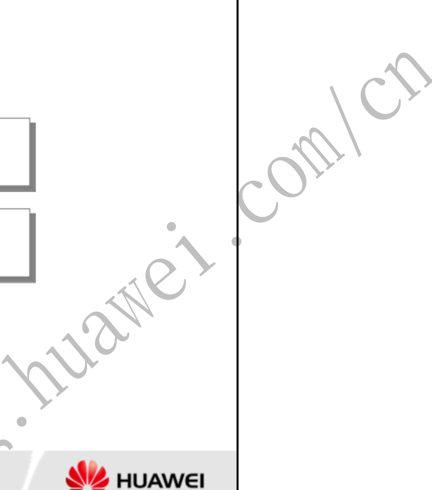
在规模较小的企业网络中，可以把所有的路由器划分到同一个区域中，同一个OSPF区域中的路由器中的LSDB是完全一致的。OSPF区域号可以手动配置，为了便于将来的网络扩展，推荐将该区域号设置为0，即骨干区域。

OSPF开销




```
[RTA]interface GigabitEthernet 0/0/0
[RTA-GigabitEthernet0/0/0]ospf cost 20
```

```
[RTB]ospf
[RTB-ospf-1]bandwidth-reference 10000
```



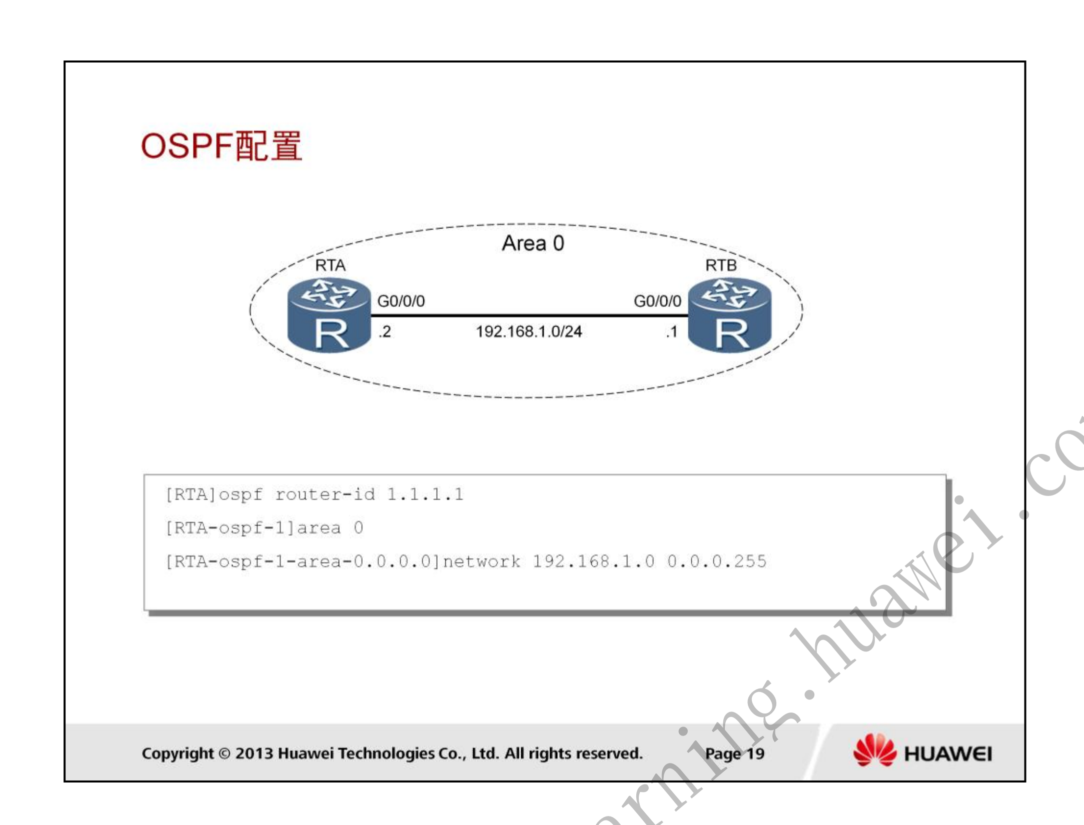
- OSPF的开销计算公式为带宽参考值/带宽。
- 可以通过bandwidth-reference命令来设置带宽参考值。

  HUAWEI

OSPF基于接口带宽计算开销，计算公式为：接口开销=带宽参考值÷带宽。带宽参考值可配置，缺省为100Mbit/s。以此，一个64kbit/s串口的开销为1562，一个E1接口（2.048 Mbit/s）的开销为48。

命令bandwidth-reference可以用来调整带宽参考值，从而可以改变接口开销，带宽参考值越大，开销越准确。在支持10Gbit/s速率的情况下，推荐将带宽参考值提高到10000Mbit/s来分别为10 Gbit/s、1 Gbit/s和100Mbit/s的链路提供1、10和100的开销。注意，配置带宽参考值时，需要在整个OSPF网络中统一进行调整。

另外，还可以通过ospf cost命令来手动为一个接口调整开销，开销值范围是1~65535，缺省值为1。



在配置OSPF时，需要首先使能OSPF进程。

命令**ospf** [*process id*]用来使能OSPF，在该命令中可以配置进程ID。如果没有配置进程ID，则使用1作为缺省进程ID。

命令**ospf** [*process id*] [**router-id** <*router-id*>]既可以使能OSPF进程，还同时可以用于配置Router ID。在该命令中，*router-id*代表路由器的ID。

命令**network**用于指定运行OSPF协议的接口，在该命令中需要指定一个反掩码。反掩码中，“0”表示此位必须严格匹配，“1”表示该地址可以为任意值。

配置验证

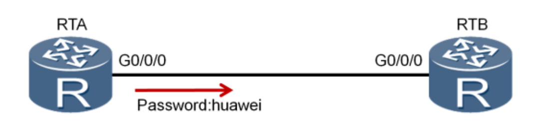
```
[RTA]display ospf peer

         OSPF Process 1 with Router ID 1.1.1.1
                 Neighbors

Area 0.0.0.0 interface 192.168.1.2(GigabitEthernet0/0/0)'s
neighbors
 Router ID: 2.2.2.2          Address: 192.168.1.1
   State: Full  Mode:Nbr is  Slave  Priority: 1
   DR: 192.168.1.2  BDR: 192.168.1.1  MTU: 0
   Dead timer due in 40  sec
   Retrans timer interval: 5
   Neighbor is up for 00:00:31
   Authentication Sequence: [ 0 ]
```


命令**display ospf peer**可以用于查看邻居相关的属性，包括区域、邻居的状态、邻接协商的主从状态以及DR和BDR情况。

OSPF认证




```
[RTA]interface GigabitEthernet0/0/0
[RTA-GigabitEthernet0/0/0]ospf authentication-mode md5 1 huawei
```

- 华为ARG3系列路由器运行OSPF时，支持两种认证方式：区域认证和接口认证。



OSPF支持简单认证及加密认证功能，加密认证对潜在的攻击行为有更强的防范性。OSPF认证可以配置在接口或区域上，配置接口认证方式的优先级高于区域认证方式。

接口或区域上都可以运行**ospf authentication-mode** { **simple** [[**plain**] <**plain-text**> | **cipher** <**cipher-text**>] | **null** } 命令来配置简单认证，参数**plain**表示使用明文密码，参数**cipher**表示使用密文密码，参数**null**表示不认证。

命令**ospf authentication-mode** { **md5** | **hmac-md5** } [*key-id* { **plain** <**plain-text** >| [**cipher**] <**cipher-text**>}] 用于配置加密认证，MD5是一种保证链路认证安全的加密算法（具体配置已在举例中给出），参数*key-id*表示接口加密认证中的认证密钥ID，它必须与对端上的*key-id*一致。

配置验证

```
<RTA>terminal debugging
<RTA>debugging ospf packet
Aug 19 2013 08:10:06.850.2+00:00 R2 RM/6/RMDEBUG:  Source Address:
192.168.1.2
Aug 19 2013 08:10:06.850.3+00:00 R2 RM/6/RMDEBUG:  Destination
Address: 224.0.0.5
……
Aug 19 2013 08:10:06.850.6+00:00 R2 RM/6/RMDEBUG:  Area: 0.0.0.0,
Chksum: 0
Aug 19 2013 08:10:06.850.7+00:00 R2 RM/6/RMDEBUG:  AuType: 02
Aug 19 2013 08:10:06.850.8+00:00 R2 RM/6/RMDEBUG:  Key(ascii): * *
* * * * * *
```

 Page 22



在启用认证功能之后，可以在终端上进行调试来查看认证过程。

debugging ospf packet命令用来指定调试OSPF报文，然后便可以查看认证过程，以确定认证配置是否成功。



总结

- OSPF Hello报文中Router Dead Interval字段的作用是什么?
- 在广播网络中，DR和BDR用来接收链路状态更新报文的地址是什么?

 Page 23 

1. Hello报文中的Router Dead Interval字段代表死亡间隔，如果在此时间内未收到邻居发来的Hello报文，则认为邻居失效。死亡间隔是Hello间隔的4倍，在广播网络上缺省为40秒（因为Hello间隔缺省为10秒）。
2. 在广播网络上，DR和BDR都使用组播地址224.0.0.6来接收链路状态更新报文。


谢谢
www.huawei.com

Module-5

网络应用的实现


DHCP原理与配置
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

在大型企业网络中，会有大量的主机或设备需要获取IP地址等网络参数。如果采用手工配置，工作量大且不好管理，如果有用户擅自修改网络参数，还有可能会造成IP地址冲突等问题。使用动态主机配置协议DHCP（Dynamic Host Configuration Protocol）来分配IP地址等网络参数，可以减少管理员的工作量，避免用户手工配置网络参数时造成的地址冲突。

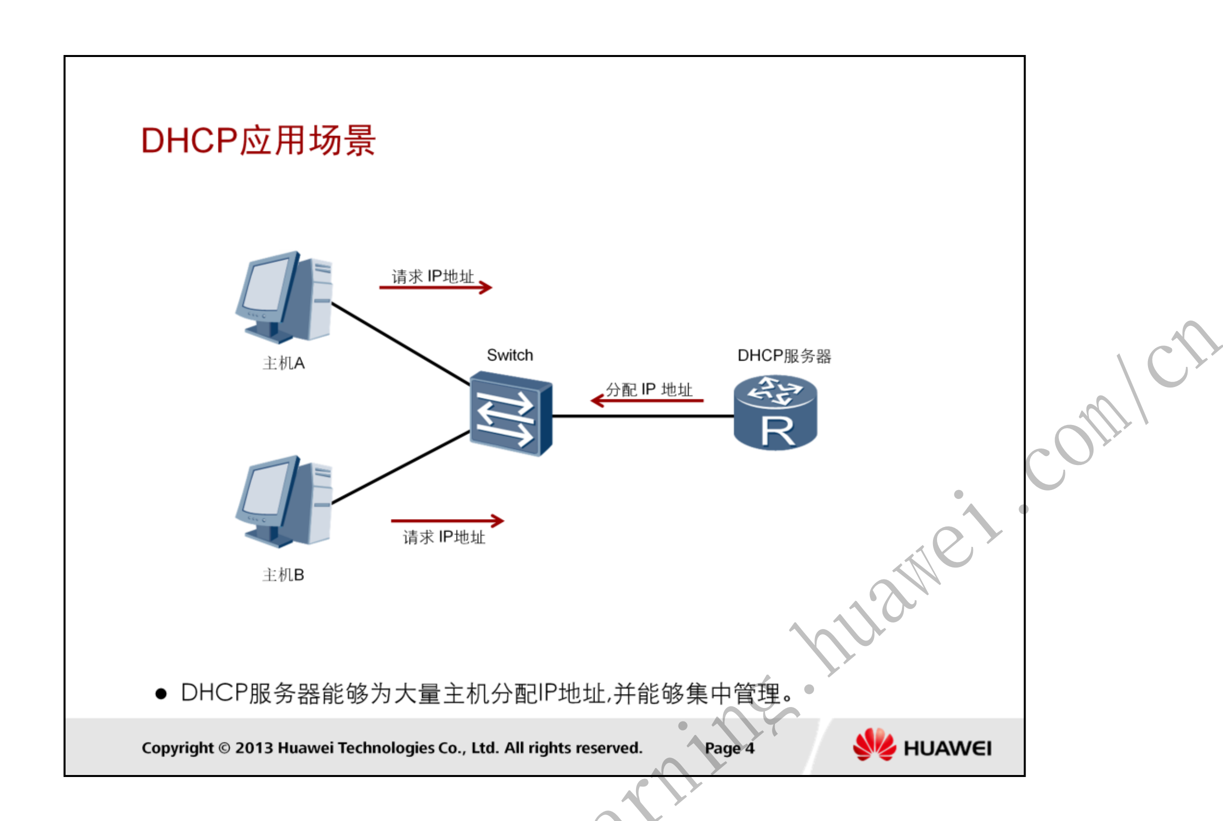




学习目标

学完本课程后，您应该能：

- 掌握DHCP的应用场景
- 掌握DHCP的基本原理
- 掌握DHCP的基本配置



在大型企业网络中，一般会有大量的主机等终端设备。每个终端都需要配置IP地址等网络参数才能接入网络。在小型网络中，终端数量很少，可以手动配置IP地址。但是在大中型网络中，终端数量很多，手动配置IP地址工作量大，而且配置时容易导致IP地址冲突等错误。

DHCP可以为网络终端动态分配IP地址，解决了手工配置IP地址时的各种问题。

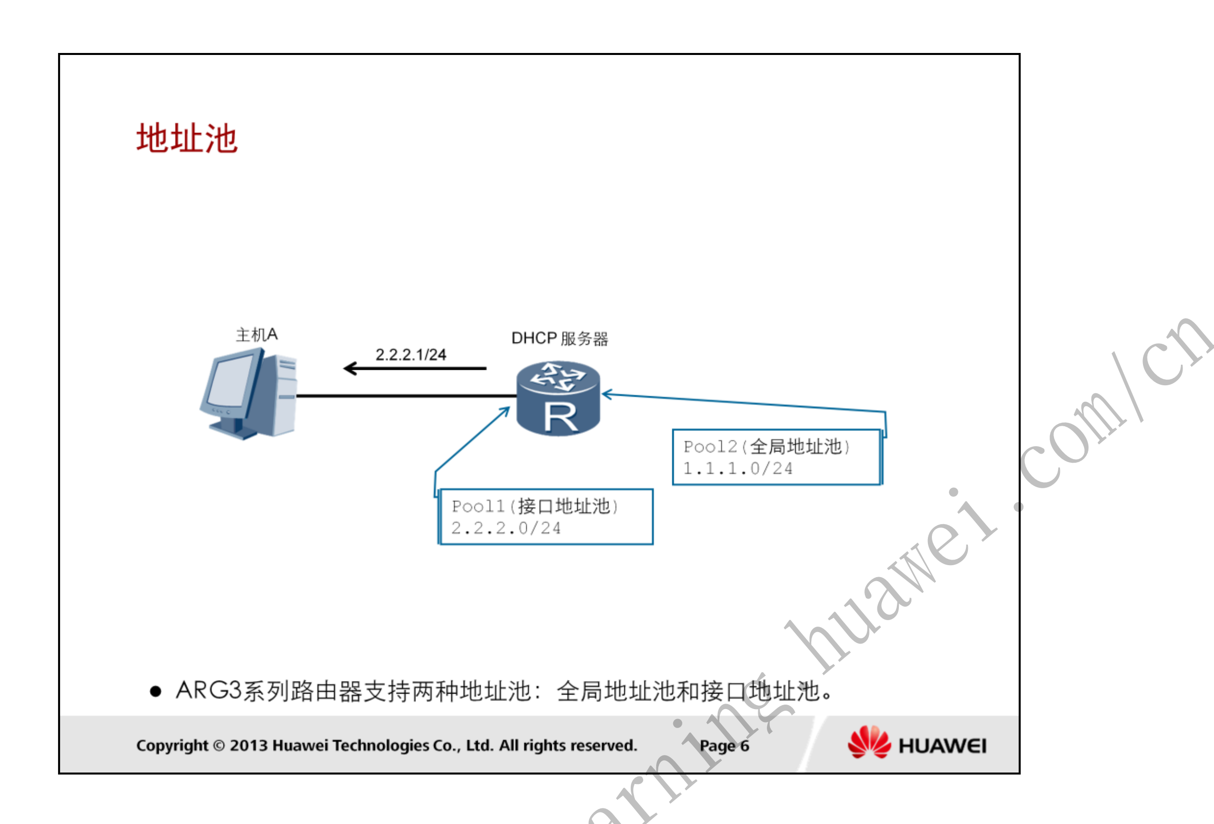
DHCP 报文类型

报文类型	含义
DHCP DISCOVER	客户端用来寻找DHCP服务器。
DHCP OFFER	DHCP服务器用来响应DHCP DISCOVER报文，此报文携带了各种配置信息。
DHCP REQUEST	客户端请求配置确认，或者续借租期。
DHCP ACK	服务器对REQUEST报文的确认响应。
DHCP NAK	服务器对REQUEST报文的拒绝响应。
DHCP RELEASE	客户端要释放地址时用来通知服务器。

 Page 5



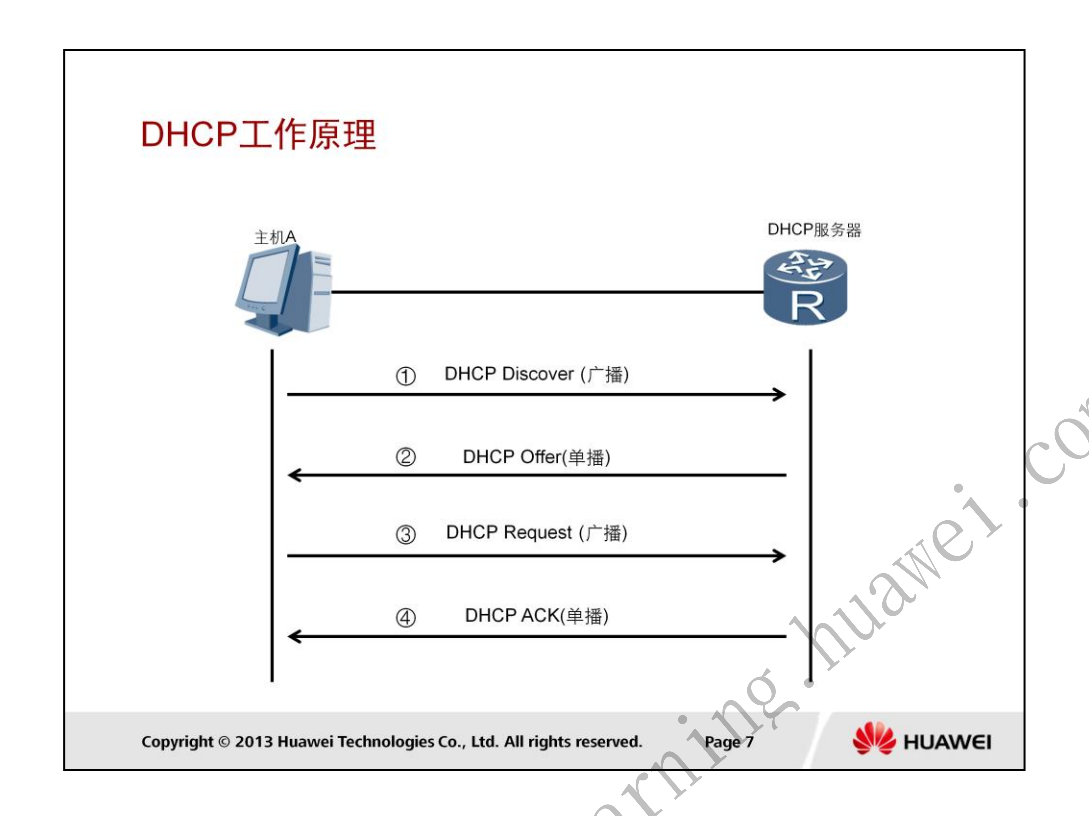
1. DHCP客户端初次接入网络时，会发送DHCP发现报文（DHCP Discover），用于查找和定位DHCP服务器。
2. DHCP服务器在收到DHCP发现报文后,发送DHCP提供报文（DHCP Offer），此报文中包含IP地址等配置信息。
3. 在DHCP客户端收到服务器发送的DHCP提供报文后，会发送DHCP请求报文（DHCP Request），另外在DHCP客户端获取IP地址并重启后，同样也会发送DHCP请求报文，用于确认分配的IP地址等配置信息。DHCP客户端获取的IP地址租期快要到期时，也发送DHCP请求报文向服务器申请延长IP地址租期。
4. 收到DHCP客户端发送的DHCP请求报文后，DHCP服务器会回复DHCP确认报文（DHCP ACK）。客户端收到DHCP确认报文后，会将获取的IP地址等信息进行配置和使用。
5. 如果DHCP服务器收到DHCP-REQUEST报文后，没有找到相应的租约记录，则发送DHCP-NAK报文作为应答，告知DHCP客户端无法分配合适IP地址。
6. DHCP客户端通过发送DHCP释放报文（DHCP Release）来释放IP地址。收到DHCP释放报文后，DHCP服务器可以把该IP地址分配给其他DHCP客户端。



ARG3系列路由器和X7系列交换机都可以作为DHCP服务器，为主机等设备分配IP地址。DHCP服务器的地址池是用来定义分配给主机的IP地址范围，有两种形式。

1. 接口地址池为连接到同一网段的主机或终端分配IP地址。可以在服务器的接口下执行dhcp select interface命令，配置DHCP服务器采用接口地址池的DHCP服务器模式为客户端分配IP地址。
2. 全局地址池为所有连接到DHCP服务器的终端分配IP地址。可以在服务器的接口下执行dhcp select global命令，配置DHCP服务器采用全局地址池的DHCP服务器模式为客户端分配IP地址。

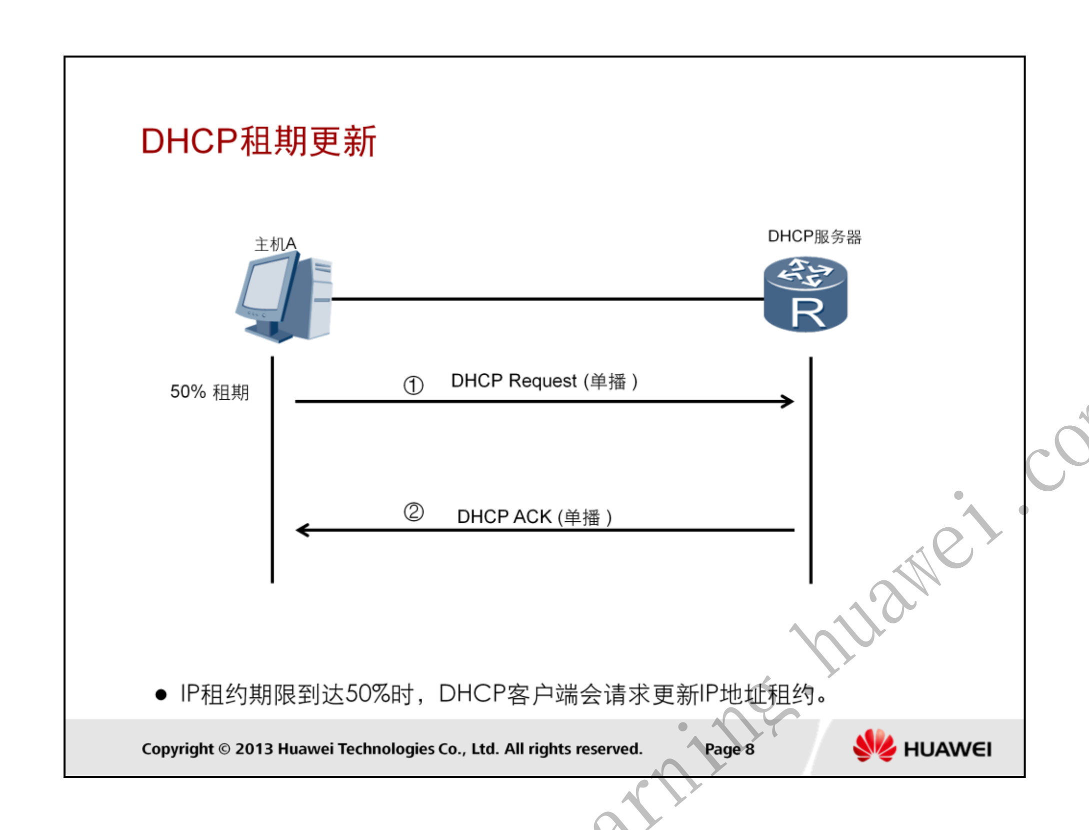
接口地址池的优先级比全局地址池高。配置了全局地址池后，如果又在接口上配置了地址池，客户端将会从接口地址池中获取IP地址。在X7系列交换机上，只能在VLANIF逻辑接口上配置接口地址池。



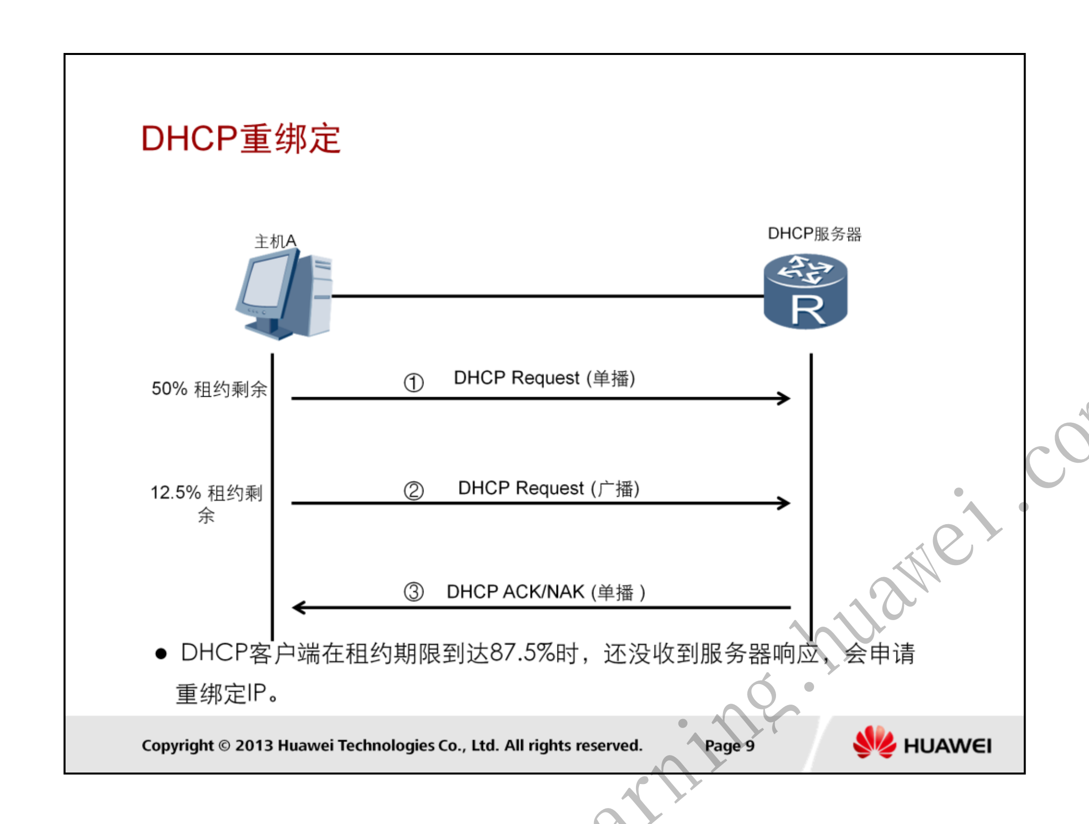
为了获取IP地址等配置信息，DHCP客户端需要和DHCP服务器进行报文交互。

首先，DHCP客户端发送DHCP发现报文来发现DHCP服务器。DHCP服务器会选取一个未分配的IP地址，向DHCP客户端发送DHCP提供报文。此报文中包含分配给客户端的IP地址和其他配置信息。如果存在多个DHCP服务器，每个DHCP服务器都会响应。

如果有多个DHCP服务器向DHCP客户端发送DHCP提供报文，DHCP客户端将会选择收到的第一个DHCP提供报文，然后发送DHCP请求报文，报文中包含请求的IP地址。收到DHCP请求报文后，提供该IP地址的DHCP服务器会向DHCP客户端发送一个DHCP确认报文，包含提供的IP地址和其他配置信息。DHCP客户端收到DHCP确认报文后，会发送免费ARP报文，检查网络中是否有其他主机使用分配的IP地址。如果指定时间内没有收到ARP应答，DHCP客户端会使用这个IP地址。如果有主机使用该IP地址，DHCP客户端会向DHCP服务器发送DHCP拒绝报文，通知服务器该IP地址已被占用。然后DHCP客户端会向服务器重新申请一个IP地址。

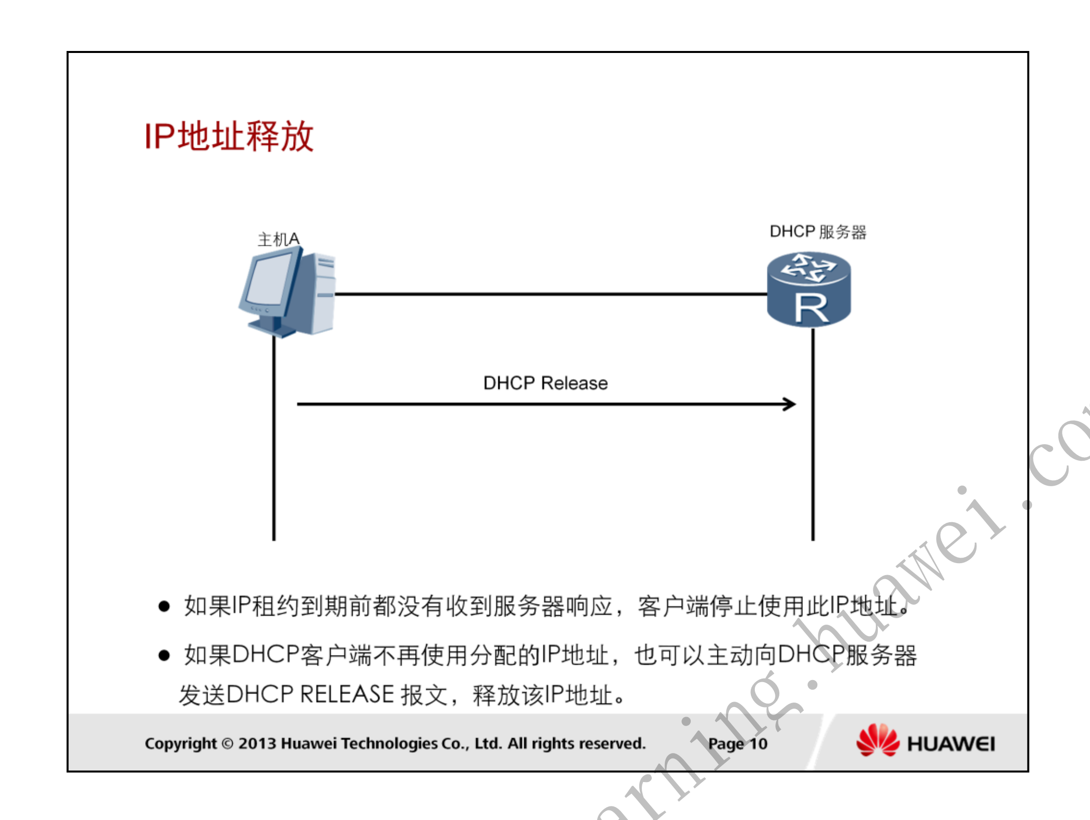


申请到IP地址后，DHCP客户端中会保存三个定时器，分别用来控制租期更新，租期重绑定和租期失效。DHCP服务器为DHCP客户端分配IP地址时会指定三个定时器的值。如果DHCP服务器没有指定定时器的值，DHCP客户端会使用缺省值，缺省租期为1天。默认情况下，还剩下50%的租期时，DHCP客户端开始租约更新过程，DHCP客户端向分配IP地址的服务器发送DHCP请求报文来申请延长IP地址的租期。DHCP服务器向客户端发送DHCP确认报文，给予DHCP客户端一个新的租期。

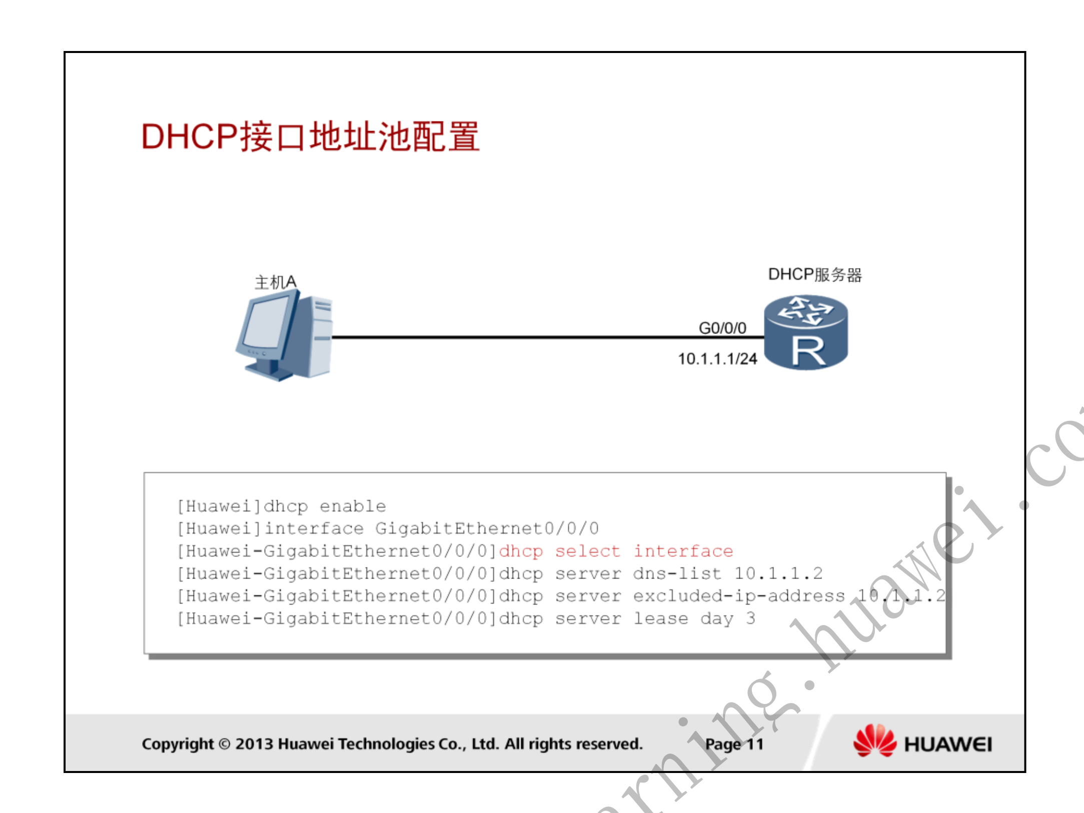


DHCP客户端发送DHCP请求报文续租时，如果DHCP客户端没有收到DHCP服务器的DHCP应答报文。默认情况下，重绑定定时器在租期剩余12.5%的时候超时，超时后，DHCP客户端会认为原DHCP服务器不可用，开始重新发送DHCP请求报文。网络上任何一台DHCP服务器都可以应答DHCP确认或DHCP非确认报文。

如果收到DHCP确认报文，DHCP客户端重新进入绑定状态，复位租期更新定时器和重绑定定时器。如果收到DHCP非确认报文，DHCP客户端进入初始化状态。此时，DHCP客户端必须立刻停止使用现有IP地址，重新申请IP地址。



租期定时器是地址失效进程中的最后一个定时器，超时时间为IP地址的租期时间。如果DHCP客户端在租期失效定时器超时前没有收到服务器的任何回应，DHCP客户端必须立刻停止使用现有IP地址，发送DHCP Release报文，并进入初始化状态。然后，DHCP客户端重新发送DHCP发现报文，申请IP地址。



DHCP支持配置两种地址池，包括全局地址池和接口地址池。

dhcp enable命令用来使能DHCP功能。在配置DHCP服务器时，必须先执行dhcp enable命令，才能配置DHCP的其他功能并生效。

dhcp select interface命令用来关联接口和接口地址池，为连接到接口的主机提供配置信息。在本示例中，接口GigabitEthernet 0/0/0被加入接口地址池中。

dhcp server dns-list命令用来指定接口地址池下的DNS服务器地址。

dhcp server excluded-ip-address命令用来配置接口地址池中不参与自动分配的IP地址范围。

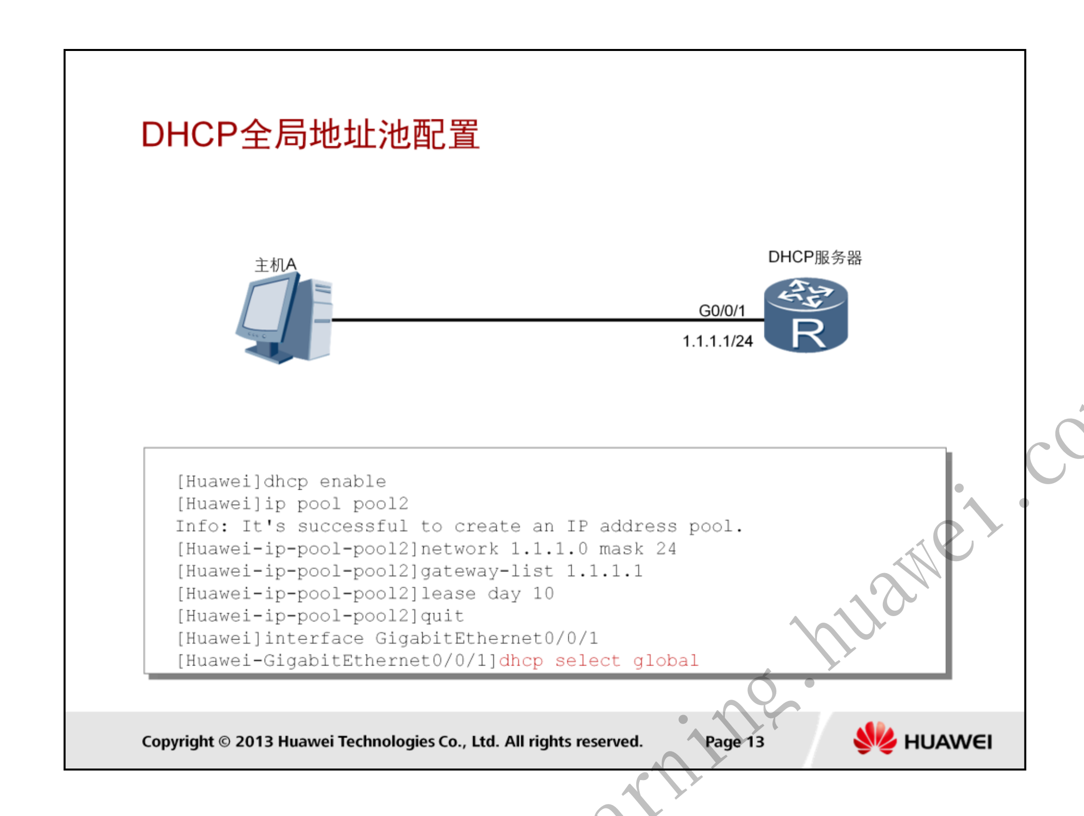
dhcp server lease命令用来配置DHCP服务器接口地址池中IP地址的租用有效期限功能。缺省情况下，接口地址池中IP地址的租用有效期限为1天。

配置验证

```
[Huawei]display ip pool
 Pool-name      : GigabitEthernet0/0/0
  Pool-No        : 0
  Position       : Interface       Status           : Unlocked
  Gateway-0      : 10.1.1.1
  Mask           : 255.255.255.0
  VPN instance   : --


  IP address Statistic
    Total       :253
    Used        :1          Idle        :252
    Expired     :0          Conflict    :0          Disable    :1
```


每个DHCP服务器可以定义一个或多个全局地址池和接口地址池。本例中执行**display ip pool**命令查看接口地址池的属性。display信息中包含地址池的IP地址范围，还包括IP网关，子网掩码等信息。



在本示例中，配置了一个DHCP全局地址池。

ip pool命令用来创建全局地址池。

network命令用来配置全局地址池下可分配的网段地址。

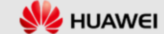
gateway-list命令用来配置DHCP服务器全局地址池的出口网关地址。

lease命令用来配置DHCP全局地址池下的地址租期。缺省情况下，IP地址租期是1天。

dhcp select global命令用来使能接口的DHCP服务器功能。

配置验证

```
[Huawei]display ip pool
 -------------------------------------------------------------
  Pool-name      : pool2
  Pool-No        : 0
  Position       : Local           Status           : Unlocked
  Gateway-0      : 1.1.1.1
  Mask           : 255.255.255.0
  VPN instance   : --
  IP address Statistic
    Total       :253
    Used        :1          Idle       :252
    Expired     :0          Conflict   :0          Disable   :0
```

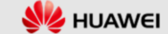
 Page 14 

display ip pool命令可以查看全局IP地址池信息。管理员可以查看地址池的网关、子网掩码、IP地址统计信息等内容，监控地址池的使用情况，了解已分配的IP地址数量，以及其他使用统计信息。



总结

- 地址池中的哪些IP地址一般会被保留?
- DHCP服务器的IP地址租期默认是多久?

 Page 15 

1. 在IP地址池中，应该排除分配给DNS等服务器的IP地址，DHCP服务器接口的IP地址等，避免IP地址冲突。
2. 默认的IP地址租期是86400秒，即一天。

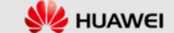

谢谢
www.huawei.com


FTP原理与配置
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

FTP是用来传送文件的协议。使用FTP实现远程文件传输的同时，还可以保证数据传输的可靠性和高效性。

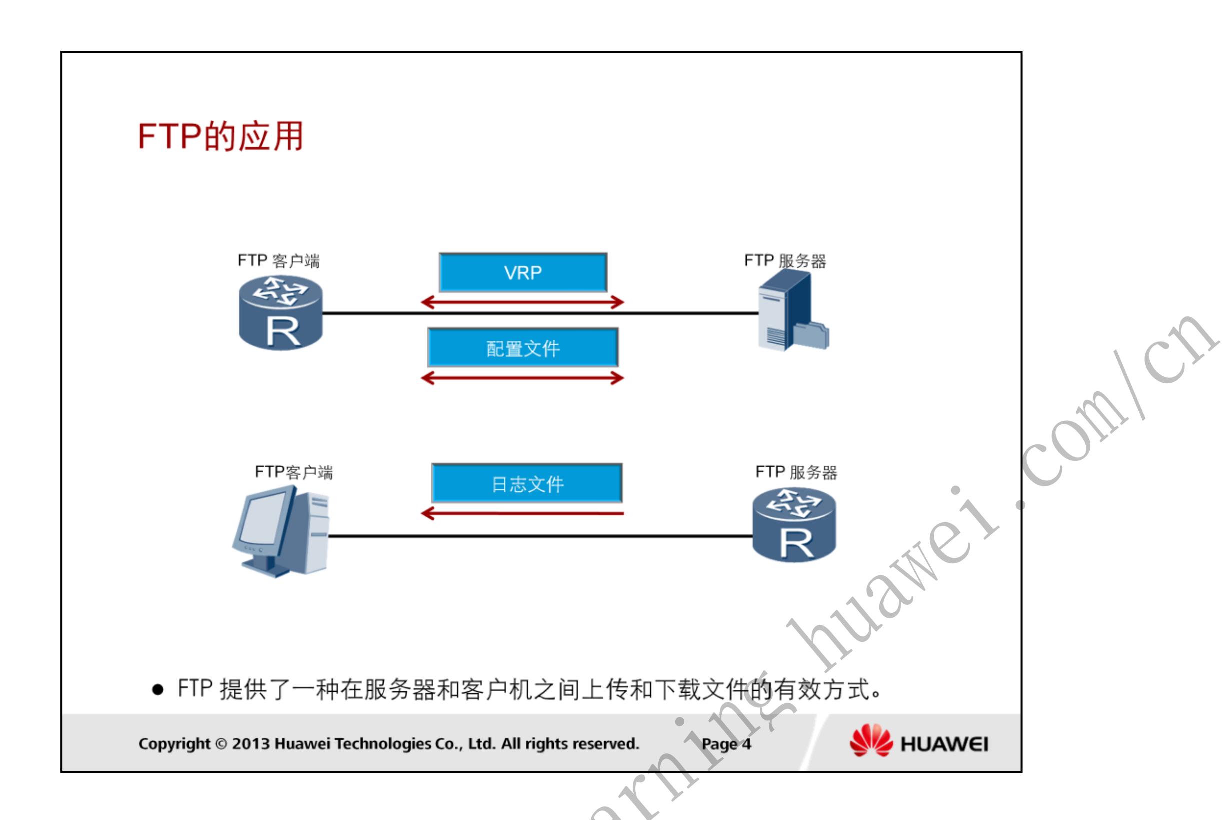




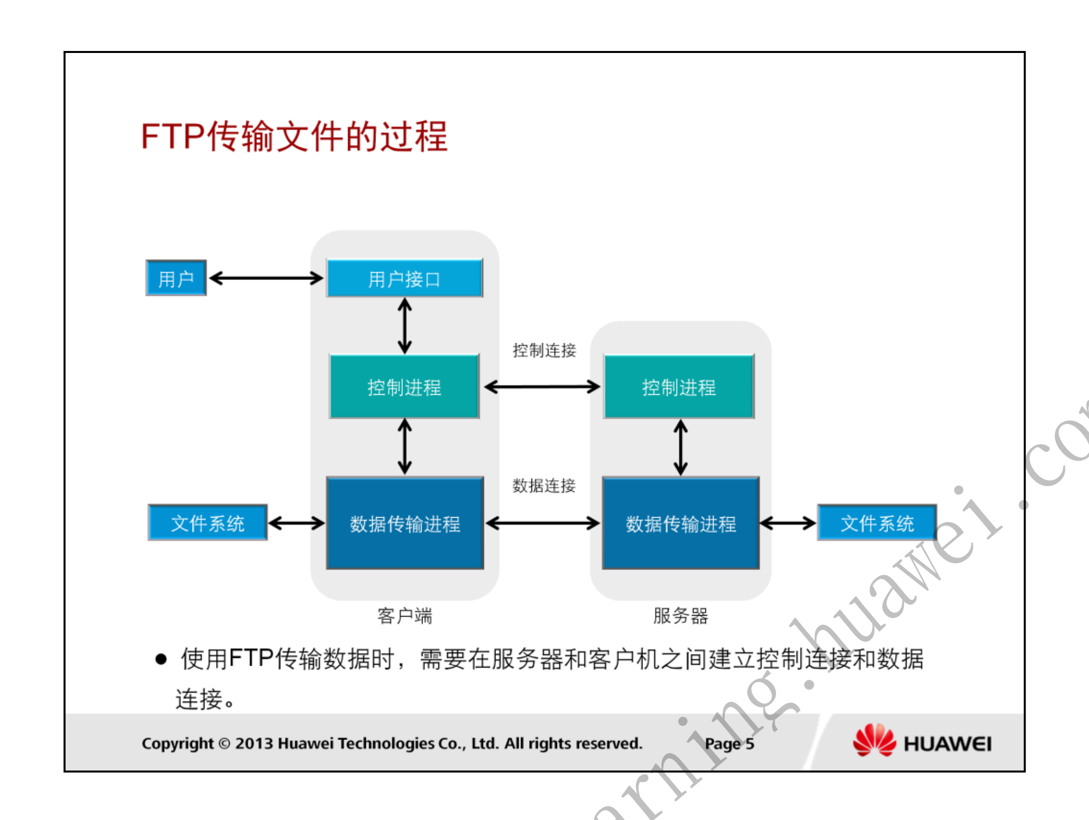
学习目标

学完本课程后，您应该能：

- 掌握FTP的工作原理
- 掌握FTP的基本配置

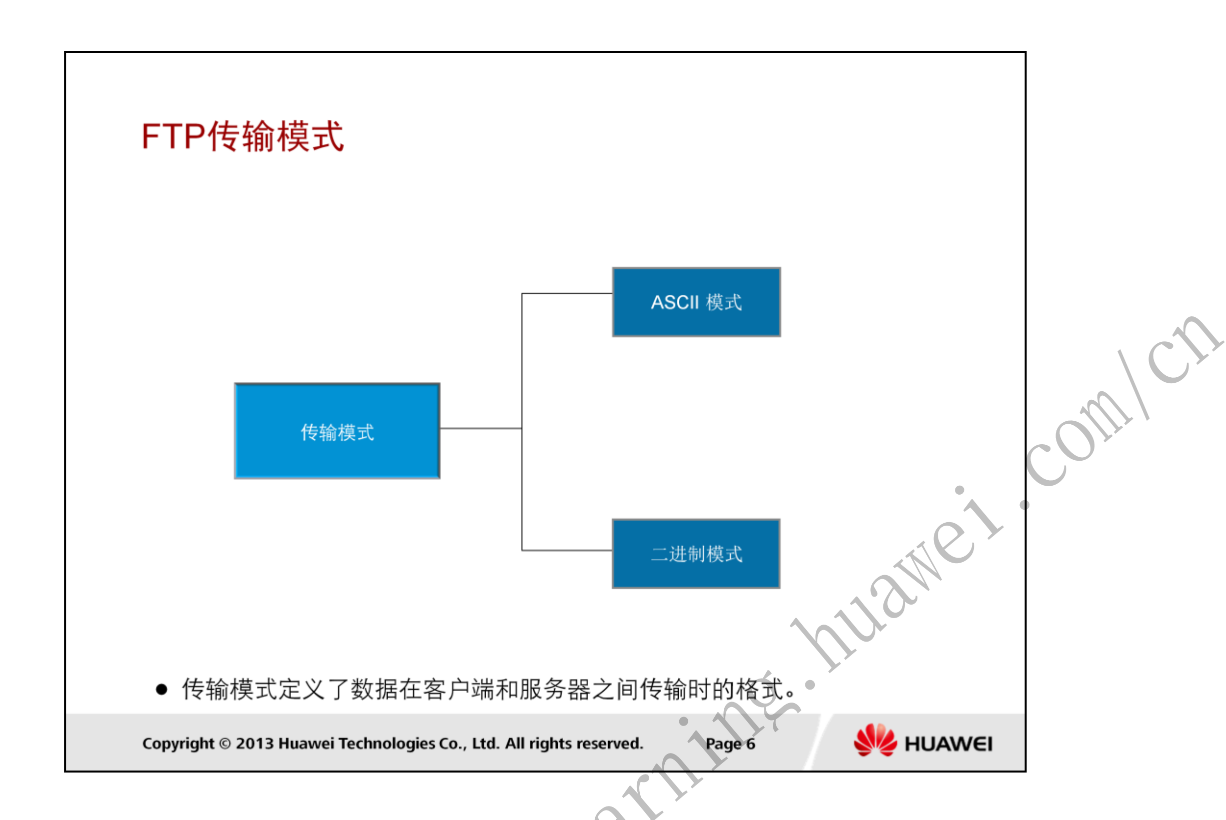


在企业网络中部署一台FTP服务器，将网络设备配置为FTP客户端，则可以使用FTP来备份或更新VRP文件和配置文件。也可以把网络设备配置为FTP服务器，将设备的日志文件保存到某台主机上方便查看。



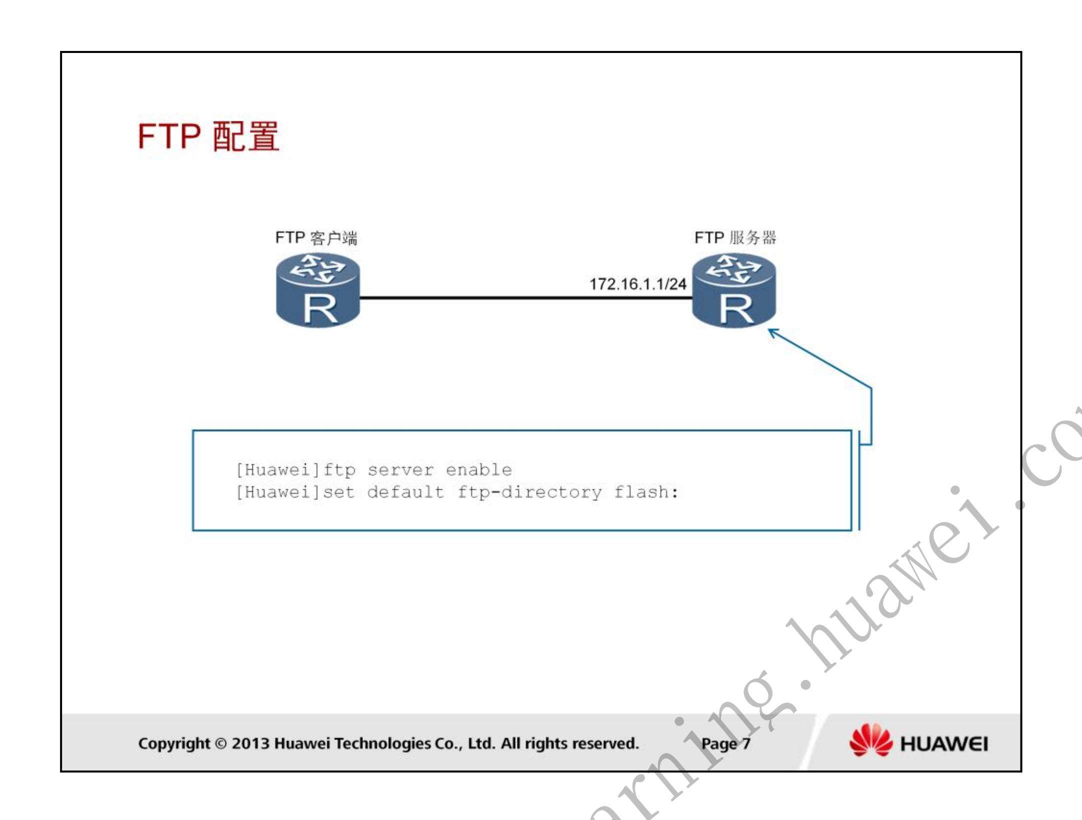
使用FTP进行文件传输时，会使用两个TCP连接。第一个连接是FTP客户端和FTP服务器间的控制连接。FTP服务器开启21号端口，等待FTP客户端发送连接请求。FTP客户端随机开启端口，向服务器发送建立连接的请求。控制连接用于在服务器和客户端之间传输控制命令。

第二个连接是FTP客户端和FTP服务器间的数据连接。服务器使用TCP的20号端口与客户端建立数据连接。通常情况下，服务器主动建立或中断数据连接。



FTP传输数据时支持两种传输模式：ASCII模式和二进制模式。

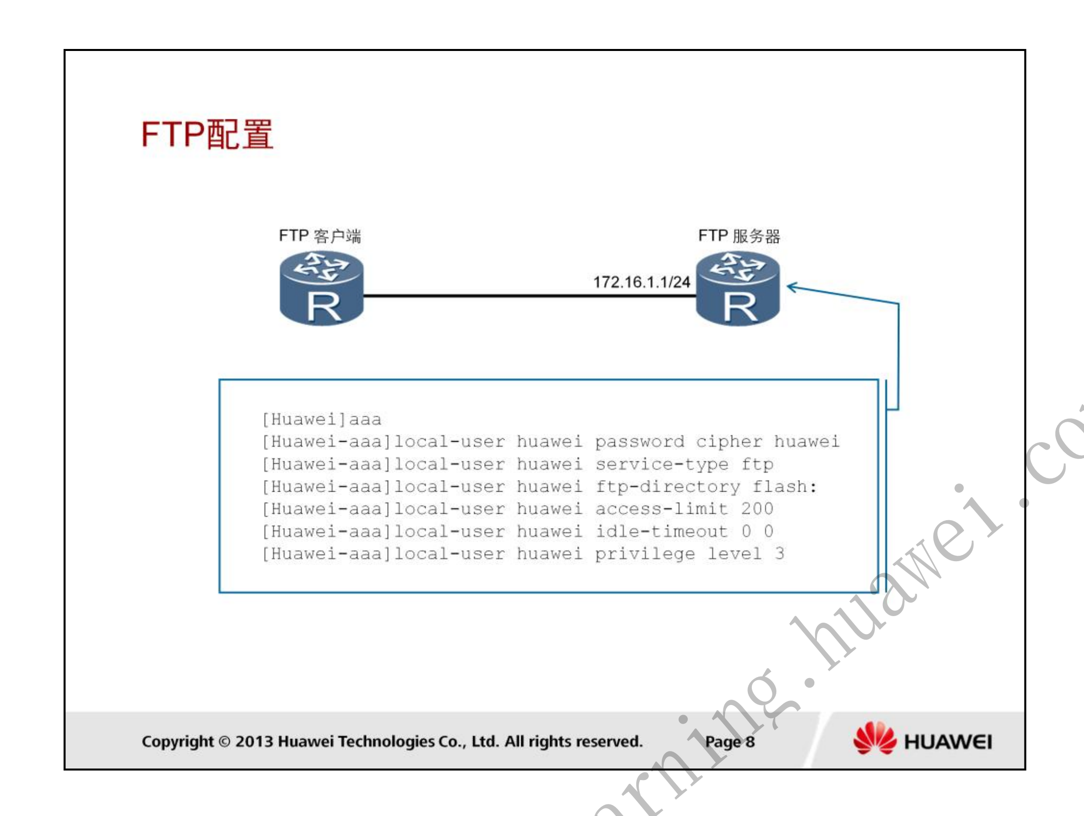
ASCII模式用于传输文本。发送端的字符在发送前被转换成ASCII码格式之后进行传输，接收端收到之后再将其转换成字符。二进制模式常用于发送图片文件和程序文件。发送端在发送这些文件时无需转换格式，即可传输。



ARG3系列路由器和X7系列交换机均可提供FTP功能。

执行**ftp server enable**命令使能FTP功能。

执行**set default ftp-directory**命令设置FTP用户的默认工作目录。



在配置FTP服务器时，可以使用AAA为每个用户分别配置登录账号和访问权限。

aaa命令用来进入AAA视图。

local-user *user-name* { **access-limit** *max-number* | **ftp-directory** *directory* | **idle-timeout** *minutes* [*seconds*] | **password** **cipher** *password* [**opt**] | **privilege level** *level* | **state** {**active** | **block** } } *命令用来创建本地用户，并配置本地用户的各项参数。

*user-name*指定用户名。

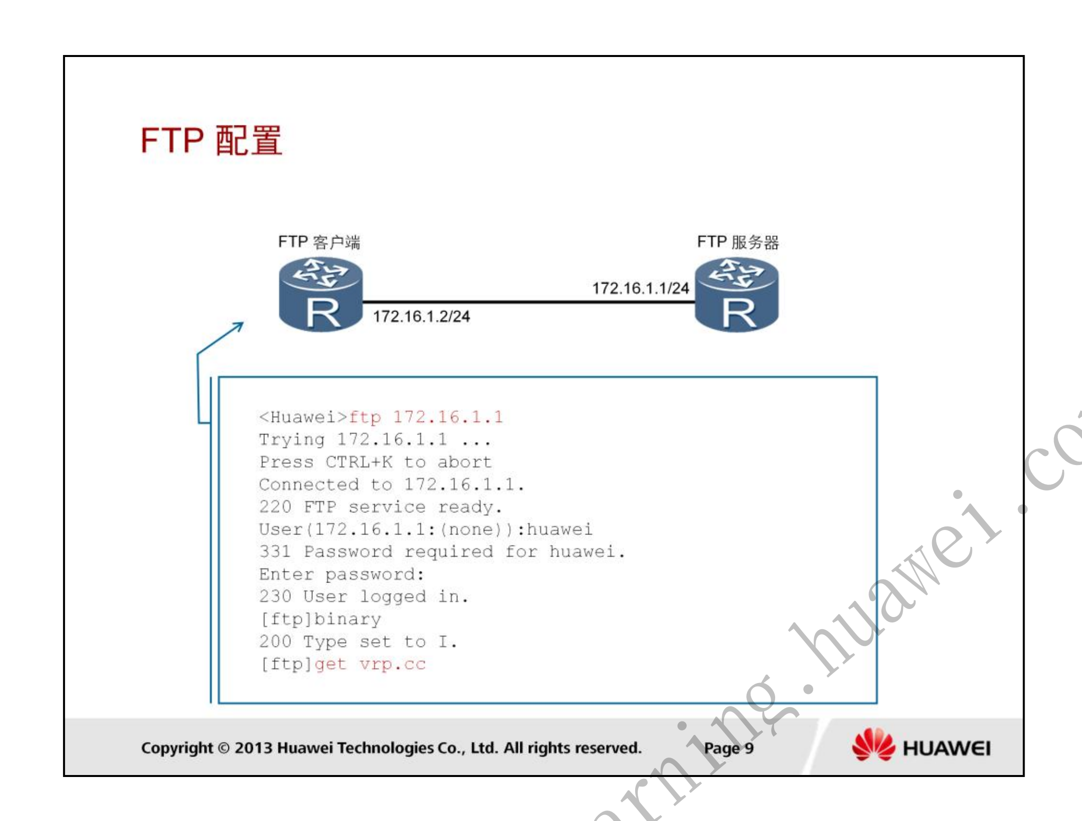
local-user huawei **service-type** ftp命令用来配置本地用户的接入类型为ftp。

ftp-directory指定FTP用户可访问的目录。如果不配置FTP用户可访问的目录，则FTP用户无法登录设备。

access-limit指定用户名可建立的最大连接数目。

idle-timeout指定用户的闲置超时时间。

privilege level指定用户的优先级。



ftp命令用来与远程FTP服务器建立控制连接，并进入FTP客户端视图。

binary命令用来在设备作为FTP客户端时设置文件传输方式为Binary模式，又称二进制模式。

缺省情况下，文件传输方式为ASCII模式。

get命令用来从远程FTP服务器下载文件并保存在本地。


总结

- FTP服务默认使用服务器哪些端口?
- 用户反馈没有权限去访问FTP服务器上的目录，应该如何解决?



1. FTP服务器需要开启TCP的21号端口来建立控制连接，20号端口来建立数据连接。
2. 如果用户无权访问任何工作目录，则需要定义一个默认的FTP目录。执行**set default ftp-directory <directory location>**命令建立默认目录。

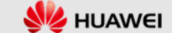

谢谢
www.huawei.com


Telnet原理与配置
HUAWEI TECHNOLOGIES CO., LTD.
HUAWEI



前言

如果企业网络中有一台或多台网络设备需要远程进行配置和管理，管理员可以使用Telnet远程连接到每一台设备上，对这些网络设备进行集中的管理和维护。

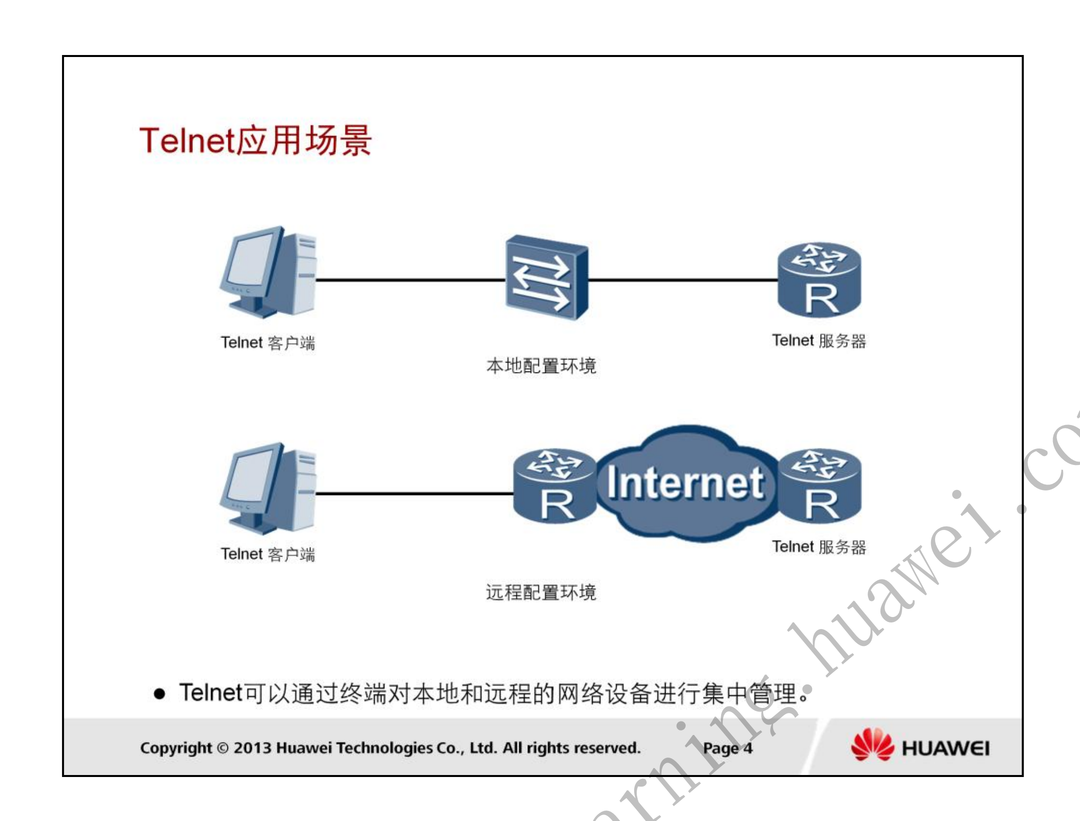




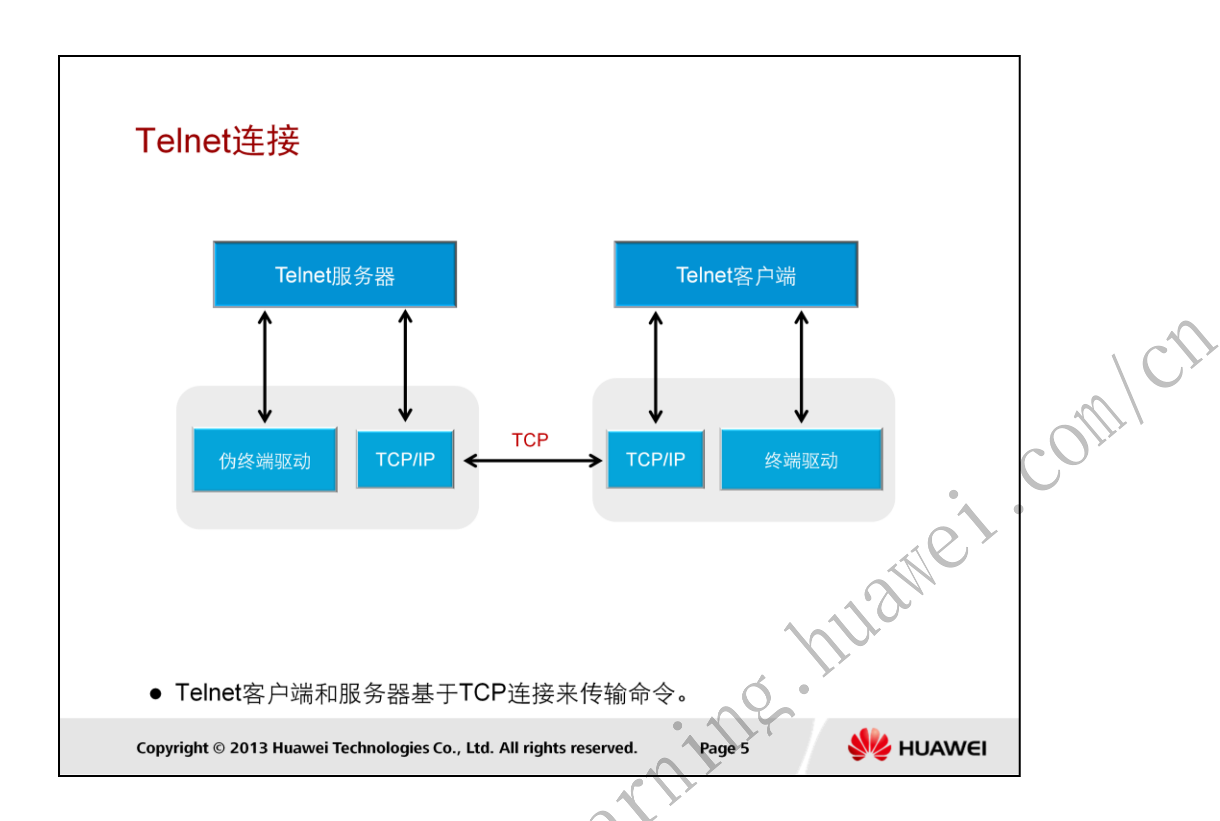
学习目标

学完本课程后，您应该能：

- 掌握Telnet的应用场景
- 掌握Telnet的工作原理
- 掌握Telnet的基本配置



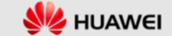
Telnet提供了一个交互式操作界面，允许终端远程登录到任何可以充当Telnet服务器的设备。Telnet用户可以像通过Console口本地登录一样对设备进行操作。远端Telnet服务器和终端之间无需直连，只需保证两者之间可以互相通信即可。通过使用Telnet，用户可以方便的实现对设备进行远程管理和维护。



Telnet以客户端/服务器模式运行。Telnet基于TCP协议，服务器端口号默认是23，服务器通过该端口与客户端建立Telnet连接。

认证模式

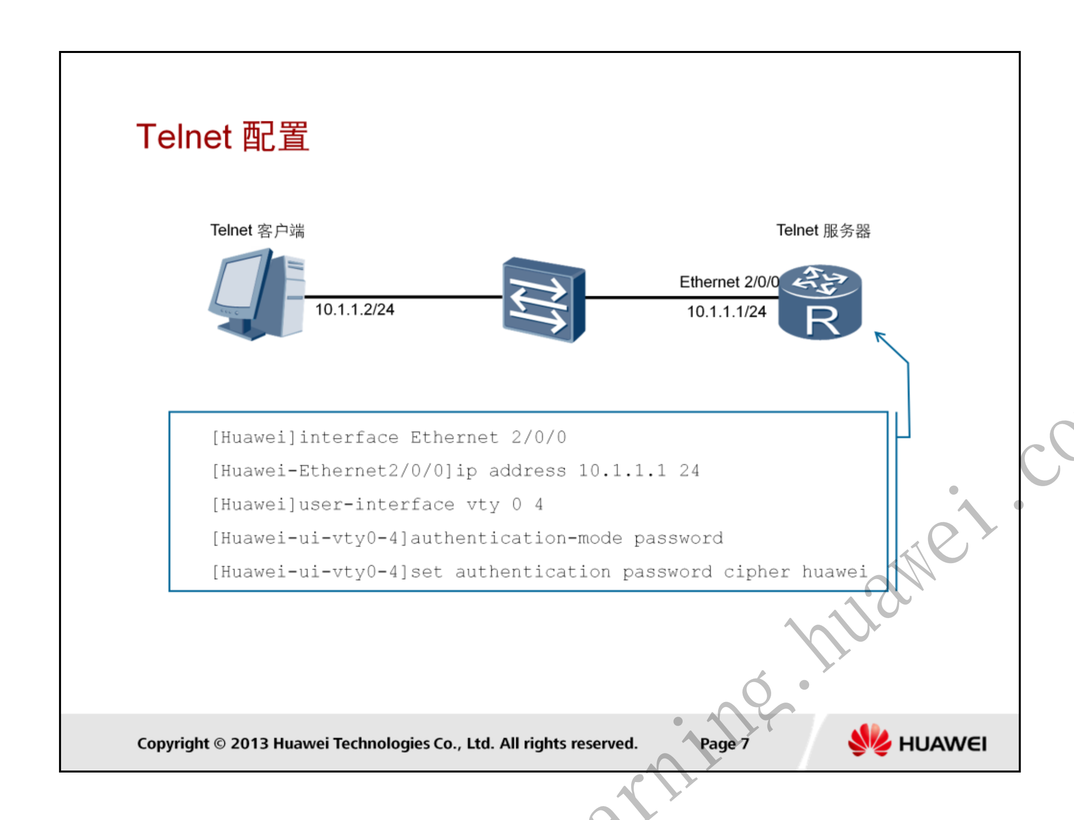
认证模式	描述
AAA	AAA 认证
Password	登录时只通过密码实现认证



在配置Telnet登录用户界面时，必须配置认证方式，否则用户无法成功登录设备。

Telnet认证有两种模式：AAA模式，密码模式。

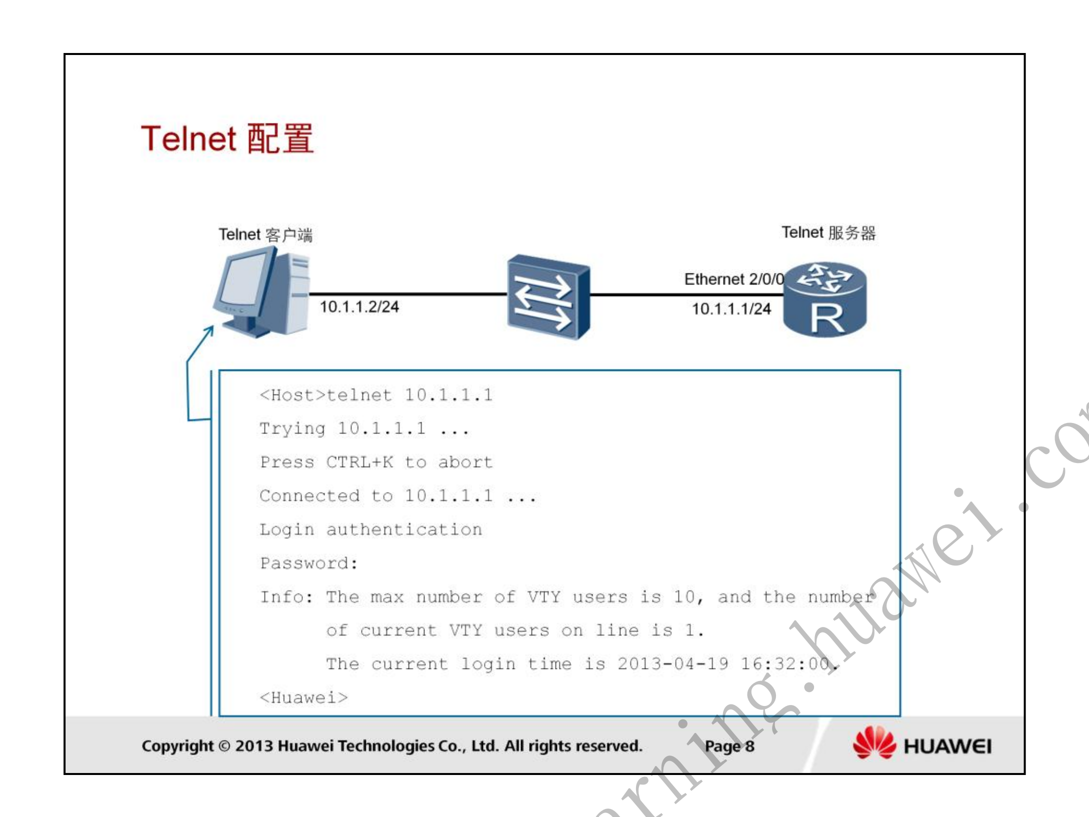
1. 当配置用户界面的认证方式为AAA时，用户登录设备时需要首先输入登录用户名和密码才能登录。
2. 当配置用户界面的认证方式为password时，用户登录设备时需要首先输入登录密码才能登录。



网络设备作为Telnet服务器，通常使用密码认证机制来认证连接到VTY接口的用户。

VTY（Virtual Type Terminal）是网络设备用来管理和监控通过Telnet方式登录的用户的界面。网络设备为每个Telnet用户分配一个VTY界面。缺省情况下，ARG3系列路由器支持的Telnet用户最大数目为5个，VTY 0 4的含义是VTY0，VTY1，VTY2，VTY3，VTY4。如果需要增加Telnet用户的登录数量，可以使用**user-interface maximum-vty**命令来调整VTY界面的数量。

执行**authentication-mode password**命令，可以配置VTY通过密码对用户进行认证。



远端设备配置为Telnet服务器之后，可以在客户端上执行**telnet**命令来与服务器建立Telnet连接。客户端会收到需要认证相关的提示信息，用户输入的认证密码需要匹配Telnet服务器上保存的密码。认证通过之后，用户就可以通过Telnet远程连接到Telnet服务器上，在本地对远端的设备进行配置和管理。



总结

- 如果网络设备已经配置完成Telnet服务，但是用户仍然不能实现远程访问，原因可能是什么？



1. 如果无法建立Telnet连接，首先验证设备是否可达。如果设备可达，再检查用户输入的密码是否正确。如果密码正确，再查看当前通过Telnet访问设备的用户数是否达到最大限制。如需增加用户数量，可以执行**user-interface maximum-vty <0-15>**命令，**0-15**表示支持的用户数。


谢谢
www.huawei.com

华为职业认证通过者权益

通过**任一项**华为职业认证，您即可在华为在线学习网站(*http://learning.huawei.com/cn*) 享有如下特权：

- 1、华为E-learning 课程学习
  - **内容**：**所有华为职业认证E-Learning课程**，扩展您在其他技术领域的技术知识
  - **方式**：请提交您的“华为账号”和注册账号的“email地址”到 *Learning@huawei.com 申请权限。*
- 2、华为培训教材下载
  - **内容**：**华为职业认证培训教材+华为产品技术培训教材**，覆盖企业网络、存储、安全等诸多领域
  - **方式**：登录*华为在线学习网站*，进入“*华为培训->面授培训*”，在具体课程页面即可下载教材。
- 3、华为在线公开课(LVC)优先参与
  - **内容**：企业网络、UC&C、安全、存储等诸多领域的职业认证课程，华为讲师授课，开班人数有限
  - **方式**：开班计划及参与方式请详见LVC排期：
    http://support.huawei.com/learning/NavigationAction!createNavi#navi[id]=_16
- 4、学习工具 eNSP
  - *eNSP (Enterprise Network Simulation Platform),* 是由华为提供的免费的、可扩展的、图形化网络仿真工具。主要对企业网路由器和交换机进行硬件模拟，完美呈现真实设备实景；同时也支持大型网络模拟，让大家在没有真实设备的情况下也能够进行实验测试。
- 另外，华为建立了知识分享平台 华为认证论坛。您可以在线与华为技术专家交流技术，与其他考生分享考试经验，一起学习华为产品技术。（http://support.huawei.com/ecommunity/bbs/list_2247.html）

